The structure and dynamics of multiplex networks

Federico Battiston

A thesis submitted to the University of London for the degree of
Doctor of Philosophy

School of Mathematical Sciences
Queen Mary University of London
United Kingdom

March 2017

Declaration

I, Federico Battiston , confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature: Date:

List of publications

Some ideas, chapters and sections of this thesis are based on a number of published manuscripts. For part I:

- "Structural measures for multiplex networks", *Physical Review E* **89** (3), 032804 (2014)
  F. Battiston, V. Nicosia, V. Latora

- "The new challenges of multiplex networks: measures and models", *The European Physical Journal Special Topics* **226** 3 401- 416 (2017)
  F. Battiston, V. Nicosia, V. Latora

- "Multilayer motif analysis of brain networks", *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27 (4), 047404 (2017)
  F. Battiston, V. Nicosia, M. Chavez, V. Latora

- "Emergence of multiplex communities in collaboration networks", *PLoS One* **11** (1) e0147451 (2016)
  F. Battiston, J. Iacovacci, V. Nicosia, G. Bianconi, V. Latora

- "Rich cores in multiplex brain networks", in preparation (2017)
  F. Battiston, F. De Vico Fallani, M. Chavez, J. Guillon, V. Latora

For part II:

- "The emergent dynamics of multiplex networks", in preparation (2017)
  F. Battiston, M. Diakonova, J. Gomez-Gardenes, V. Nicosia, V. Latora

- "Efficient exploration of multiplex networks", *New Journal of Physics* **18** 4 043035 (2017)
  F. Battiston, V. Nicosia, V. Latora

- "Interplay between consensus and coherence in a model of interacting opinions", *Physica D: Nonlinear Phenomena*, **323** 12-19 (2016)
  F. Battiston, A. Cairoli, V. Nicosia, A. Baule, V. Latora

- "Layered social influence promotes multiculturality in the Axelrod model", *Scientific Reports* **7** 1809 (2017)

F. Battiston, V. Nicosia, V. Latora, M. San Miguel

- "Determinants of public cooperation in multiplex networks", *New Journal of Physics* **19** 071002 (2017)
  F. Battiston, M. Perc, V. Latora

Abstract

Network science has provided useful answers to research questions in many fields, from biology to social science, from ecology to urban science. The first analyses of networked systems focused on binary networks, where only the topology of the connections were considered. Soon network scientists started considering weighted networks, to represent interactions with different strength, cost, or distance in space and time. Also, connections are not fixed but change over time. This is why in more recent years, a lot of attention has been devoted to temporal or time-varying networks.

We now entered the era of multi-layer networks, or multiplex networks, relational systems whose units are connected by different relationships, with links of distinct types embedded in different layers. Multiplexity has been observed in many contexts, from social network analysis to economics, medicine and ecology. The new challenge consists in applying the new tools of multiplex theory to unveil the richness associated to this novel level of complexity. How do agents organise their interactions across layers? How does this affect the dynamics of the system?

In the first part of the thesis, we provide a mathematical framework to deal with multiplex networks. We suggest metrics to unveil multiplexity from basic node, layer and edge properties to more complicated structure at the micro- and meso-scale, such as motifs, communities and cores. Measures are validated through the analysis of real-world systems such as social and collaboration networks, transportation systems and the human brain.

In the second part of the thesis we focus on dynamical processes taking place on top of multiplex networks, namely biased random walks, opinion dynamics, cultural dynamics and evolutionary game theory. All these examples show how multiplexity is crucial to determine the emergence of unexpected and instrinsically multiplex collective behavior, opening novel perspectives for the field of non-linear dynamics on networks.

Acknowledgments

As many things in life are about sharing and shared experiences, this thesis is no exception. This manuscript was directly or indirectly affected by the interactions, help and support of many people, and I will take advantage of this page to thank at least some of them. Firstly, I would like to thank Vito for supervising me during these years through a process preferentially made of calls late in the evening, couch sessions of shared writing, lots of overrated tea, blackboard drawings, whatsapp messages (better if at night) and Christmas parties. I feel I learnt a lot of things, but what matters the most is that I saw how to do this work with joy, ambition and respect, and I will try to bring this with me in the future. I am deeply indebted to Enzo for the everlasting scientific discussions, for making me a better (but still improvable) programmer, for even more overrated tea, and for being my shadow supervisor when it was needed. I am grateful to Fabrizio and Mario for hosting me in Paris, making me feel like at home from day one and at the same time showing me a different way to work and another side of science. I wish to thank all my other co-authors Adrian, Ginestra, Jeremy, Jesus, Marina, Matjaz, Maxi and in particular Andrea and Jacopo for months of working side by side. I am also grateful to Andrea and MariAngeles to act as examiners for this thesis. Eternal gratitude goes to the European Union who decided to finance the project LASAGNE making me travel all over Europe providing me with an inestimable opportunity at such a young age.

I would like to thank my family, my parents Carlotta and Marco and my sister Alice, for giving me the freedom to do what I want and supporting me along the way. It has at some times been a difficult relationship, but I wish to thank the School of Mathematical Sciences of Queen Mary University of London for these years, room 202, the first common room, Paul for making everything work, the awful food at the Curve and the hospital, the temporary building with its foxes, Tassimo, diggers and many surprises. Special thanks to my office mates in London: the italian crew Jacopo, Andrea, Moreno, Valerio, Massimo and lately Andrea with whom I shared countless coffees and many nights out, my international friends Paul, Sol, Andre', Nils, Sune, Louise, Trevor, Chao, Owen, all the others and in particular Tom for continuously hurting his knee by playing amazing squash matches with me, and Lucas who is not an office mate but still a friend. During my PhD I had the chance to attend many conferences, the complex systems school in Santa Fe, many warm-ups and workshops for young researchers where I met many colleagues and quickly turned them into great friends. I can not mention everybody so I will limit myself to the Yologozza crew, Ramon Pignatelli, Telepizza, Alberto, Elisa, Eugenio and

Jacopo. In my years in London I had the pleasure to live with many great flatmates, such as Lorenzo, Tarig, Elaine, Eirini and Valeria, in wonderful houses. I would like to thank them all, along with McCall House, the garden and the apple trees in Holloway and Crondall Street in Hoxton. I am grateful to have met Marcello, Camilla and Laura and to have shared many dinners at their place, first in Broadway Market and then in Homerton. Sincere gratitude also goes to Franco Manca for feeding me on many Friday nights, to all my favourite places in North-East London such as Gallipoli and their special vibe, the disgusting ales and the very average lagers found in the UK which make the British proud for no reason whatsoever. A big thanks to the Aramis Lab at the Institut du Cerveau et de la Moelle Epiniere in Paris and all the great people I met there. Two high fives Revos-style to Revolution'Air and all the Ultimate guys, in particular l'Equipe 3, Porte de la Chapelle, Porte de Clignancourt, Japy and Clissold Park. I thank all my friends from Rome, those from school and childhood, in particular Dario (who was in Paris with me), Jacopo (who was in London with me), Luigi, Francesco, Federica, Giulia, Silvia, Lorenzo, Annalisa, Chiara e Chiara, and all my friends from La Sapienza I still hang out with, such as Enrico, Gianluca, Alice, Nicandro, Nicolo', Camilla, Jared, Giulia, Marco, Gian Marco and Biko. I wish I could also thank in person Michele, who first taught me Python four years ago. A "super-giga" special thanks to Mimosa for sharing an amazing journey with me. At last, I wish to thank my random encounter with the book Nexus by Mark Buchanan, who totally by chance brought me into network science as a Physics Master student and possibly saved me from many years of boredom.

Table of Contents

Introduction

Unveiling a new level of complexity

The human brain, a social system and the world trade web are all examples of complex systems: they are all made by a large number of single units interacting in such a way that the global behavior is not a simple combination of the behaviors of the single components. Despite the lack of centralised control, they are characterised by the emergence of collective behavior and sometimes extreme or unpredictable events, no matter how good the knowledge of the properties of the single units is. Examples include epileptic seizures, the diffusion of online viral contents online, traffic jams and financial bubbles. Neurons are different from individuals and countries: this explains why such diverse agents have become the object of study of neuroscience, sociology and economics respectively. Yet, in all these systems reducing the problem to the analysis of the single pieces is not enough. In order to grasp many critical features observed empirically is necessary to consider the complex network of interactions connecting the different units and representing the backbone of the complex system.

Network science has provided useful answers to research questions in many fields, from biology to social science, from ecology to urban science [1–4]. The first analyses of networked systems focused on binary networks, where only the topology of the connections was taken into account [5, 6]. Soon network scientists started considering weighted networks [7], to represent interactions with different strength, cost, or distance in space and time. Also, connections are not fixed but change over time. This is why in more recent years, a lot of attention has been devoted to temporal or time-varying networks [8].

We now entered the era of *multi-layer networks*, or *multiplex networks*, relational systems whose units are connected by different relationships, with links of distinct types embedded in different layers [9–15]. *Multiplexity* [16] has been observed in many contexts, from social network analysis [17] to economics [18], medicine [19] and ecology [20].

For instance, multimodal transportation systems are now ubiquitous in large metropolitan areas, where locations can be reached by bus, underground, suburban rail and naval routes [21]. Some of the best well-known network datasets, such as the co-authorship network of scientists, the co-starring network of movie actors, or the neural system of the worm C. Elegans are multiplex networks: scientific collaborations can be distinguished according to the area of research; movies classified in several genres; neurons communicate by means of synapses or gap junctions [22].

The new challenge consists in applying the new tools of multiplex theory to unveil the richness associated to this novel level of complexity. How do agents organise their interactions across layers? How does this affect the dynamics of the system? In such sense, the research on multi-layer networks presented in this thesis is at the heart of complexity science, focusing on fundamental aspects and providing new tools for the analysis of the structure and dynamics of real-world systems.

Beyond the buzzword: when 'multiplex' really matters?

Multiplexity does not come for free: multi-layer networks require more memory and more complicated algorithms for their analysis compared to single-layer graphs. As a consequence, keeping track of the different nature of the links should be the result of a cost-benefit analysis based on a trade-off between amount of information and resources needed to store it [23].

Layered interactions have been shown to give rise to novel criticality for a wide set of dynamical processes. To name just few of them, multi-layer networks are characterised by a first-order percolation transition and are more fragile to attacks than single-layer graphs [24], are endowed with a super-diffusive regime faster than the diffusion at any single layer [25], promote novel types of Turing patterns in reaction-diffusion processes [26], novel stable synchronised states [27] and qualitatively different spreading phenomena [28].

Multi-layer networks also allow to model intertwined dynamical processes on interacting layers, as in the case of random walk and kuramoto process coupled together to model energy transport and neural dynamics in the brain [29]. Other interesting cases are represented by the diffusion of awareness on disease spreading, where informed individuals become immunised to the spreading pathogen [30], the impact of trust, modelled as a cooperation game, on information diffusion [31], and the interplay between choices based on strategy and imitation [32].

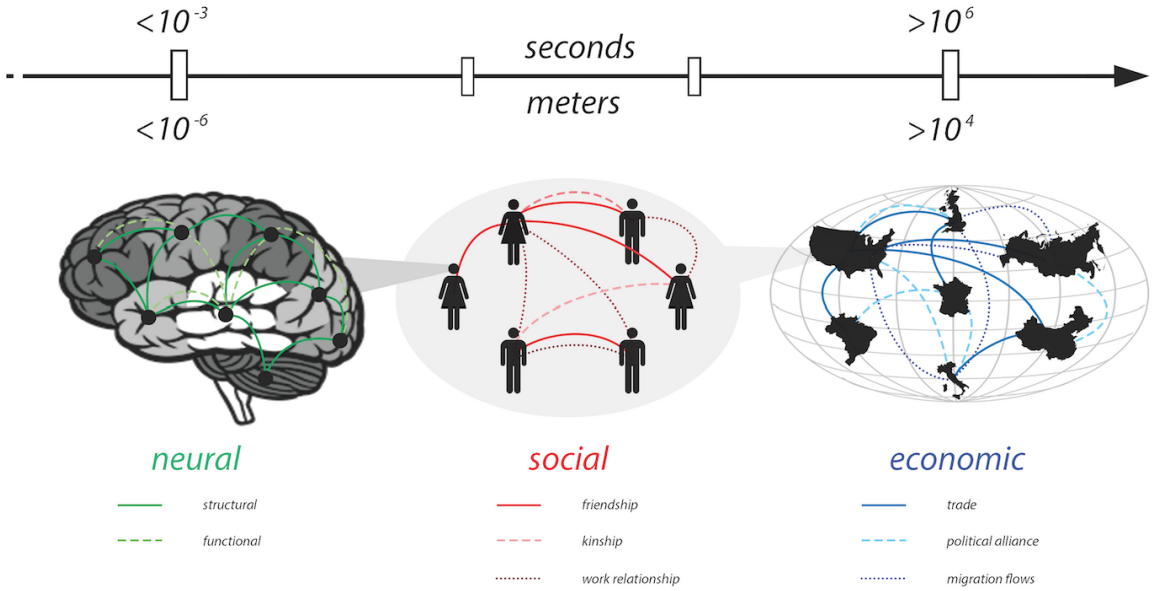


Figure 0.1: Multiplexity pervades many relational systems and may occur at very different spatio-temporal scales. Let us focus on the human being, one of the most complex systems. At the scale of microns, our brain is composed of hundreds of billions of neurons coupled through diverse actions lasting milliseconds and which work together to make us move and take decisions (neural scale). The second scale is that of the interactions among individuals, at distance of meters and ranging from seconds to weeks, regulated by a variety of social mechanisms such as social influence, homophily, norms and conventions (social scale). The largest one is that of countries, aggregates made of a large number of individuals, bond together by an intricate pattern of trade relations, political alliances and trasportation links, thousands of kilometres away and which evolve over periods of months or years (economic scale).

As not every system made of many components is complex, multiplexity is not interesting per se, but only when it produces novel behavior which can not be obtained on any layer individually, or by merging all links into an aggregated network [33].

Multiplex networks: structure and dynamics. In the first part of these thesis we measure and model the *structure* of several genuinely real-world multiplex networks, such as social networks, collaboration networks, transportation networks and human brain networks, at the micro-, meso- and macro-scale. In the second part of this thesis we investigate *dynamics* on multiplex networks, with a special focus on random walks, opinion dynamics, cultural dynamics and evolutionary dynamics of group interactions.

Part I

Structure of multiplex networks

In the first part of this thesis we propose a general formalism and a comprehensive approach to analyse the *structure* of *multiplex networks* across different scales, from basic node and edge properties to more complicated mesoscopic structures.

Single-layer networks. Let us consider a complex system made by many interacting units its basic units. By taking advantage of graph theory, in network science the N elements of the system are usually known as *nodes* (or vertices), connected through *links* (or edges) representing their pairwise interactions. This information is typically stored into an $N \times N$ adjacency matrix $\boldsymbol{A} = \{a_{ij}\}$, where $a_{ij} = 1$ if nodes i and j are connected, and $a_{ij} = 0$ otherwise. In the simplest case, each relationship is bidirectional and $\boldsymbol{A}$ is a symmetric matrix, i.e. $a_{ij} = a_{ji}$. However, such framework can be easily extended to the case of *directed* interactions, where a_{ij} describes the existence of a link from i to j and $\boldsymbol{A}$ is in general an asymmetric matrix. Relationships can have different intensities. For this reason, a network is often described in terms of a *weighted* adjacency matrix $\boldsymbol{W} = \{w_{ij}\}$, where w_{ij} is typically a non-negative number describing the strength of the connection between i and j. Negative entries are allowed as well in the case of *signed* networks, such as social networks or functional networks, where links can be either positive or negative. Single-layer networks have been widely covered in the literature, for instance in Ref. [1–4].

Networks with multiple layers of interactions. When it is possible to distinguish the nature of the ties, we are dealing with a *multiplex* network $\mathcal{M}$. Let us try to define this notion more formally [9–11]. We consider a system composed of N nodes interacting through links of M different type. An effective approach consists of embedding all edges of different types in different *layers* of interactions. The pattern of connections at layer α, with $\alpha = 1, \dots, M$, can be described by an adjacency matrix $A^{[\alpha]} = \{a_{ij}^{[\alpha]}\}$, where the entry $a_{ij}^{[\alpha]}$ for $i, j = 1, 2, \dots, N$ is either 1 or 0, depending on the existence or absence of a link from node i to node j at layer α. Notice that, in the notation we will use for the rest of the thesis, subscripts are given in roman letters and indicate nodes, while superscripts are given in greek letters and indicate layers. Hence all the *intra-layer* connections of our system are encoded in the set $\boldsymbol{A}$ of M adjacency matrices $\boldsymbol{A} = \{A^{[1]}, A^{[2]}, \dots, A^{[M]}\}$, with $\boldsymbol{A} \in \mathbb{R}_{\geq 0}^{M \times N \times N}$ [15].

In some systems there is no explicit cost to pay associated to changing layer, i.e. to switching from the use of links of layer α to links of layer β at a generic node i. In this case the multiplex network is fully-defined by specifying $\boldsymbol{A}$, i.e.:

$$\mathcal{M} \equiv \boldsymbol{A} = \{A^{[1]}, A^{[2]}, \dots, A^{[M]}\} \tag{1}$$

Similar structures have also been studied in the mathematical literature under the name

of edge-colored graphs [35]. We note here though that in multiplex networks links of different types, or colors, between the same units can coexist. The formalism can be easily extended to the case of weighted multiplex networks $\boldsymbol{W} = \{W^{[1]}, W^{[2]}, \ldots, W^{[M]}\}$ [15, 34], with directed or signed links.

However, there are cases of practical importance where switching between operating in different modes comes at a cost. Such is the case with multimodal urban transportation networks, where the main locations in a city (naturally represented as the N nodes of a multiplex network) include, for instance, bike, bus, underground and train stations (representing the $M = 4$ modes seen as layers), and it can take a varying amount of time to move from one transportation medium to another at a given location. These real-life systems can be modelled by a multiplex framework with a defined weight of *inter-layer* links, so that the cost associated with switching layer could be taken into account and compared with the cost to move between nodes in the same layer. Thus, for each node i, the weight of interactions between two different layers α and β (or some similar information about the inter-layer connectivity) is given by the $M \times M$ coupling matrix $C_i = \{c_i^{[\alpha\beta]}\}$, where the entry $c_i^{[\alpha\beta]}$ for $\alpha, \beta = 1, 2, \ldots, M$ is either 1 or 0, depending on whether or not it is possible to go from layer α to layer β at node i, or, more generally, a non-negative real number describing the weight of the inter-layer link at node i from layer α to layer β. These inter-layer connections are stored in the set of N coupling matrices, one for each node, $\boldsymbol{C} = \{C_1, C_2, \ldots, C_N\}$, with $\boldsymbol{C} \in \mathbb{R}_{\geq 0}^{N \times M \times M}$. Thus, a general multiplex network $\mathcal{M}$ is described by the two sets $\boldsymbol{A}$ and $\boldsymbol{C}$:

$$\mathcal{M} \equiv (\boldsymbol{A}, \boldsymbol{C}) = \{A^{[1]}, A^{[2]}, \ldots, A^{[M]}, C_1, C_2, \ldots, C_N\} \tag{2}$$

where $\boldsymbol{A}$ and $\boldsymbol{C}$ account, respectively, for the intra-layer and inter-layer connectivity. These multiplex networks are formally defined by giving M $N \times N$ adjacency matrices that describe the connections between nodes at each of the layers, as well as N $M \times M$ coupling matrices describing the connections between layers at each of the nodes.

In some particular cases, the coupling between two layers is a node-independent property, and $\boldsymbol{C}$ is described by a single matrix $C_i = C \ \forall i$, with $C = \{c^{[\alpha\beta]}\}$. Moreover, to simplify the treatment (especially analytically) it is sometimes assumed that all the layers are coupled with the same intensity, i.e. $c^{[\alpha\beta]} = c > 0 \ \forall \ \alpha, \beta$, so that the form of $\boldsymbol{C}$ further reduces to a single positive number c controlling the weight of the inter-layer links with respect to the weights of the intra-layer links.

A particular case: temporal networks. In other cases the structure of $\boldsymbol{C}$ simplifies for different reasons, since not all layers of a multiplex system are necessarily coupled directly. This is reflected in the number and positions of zeros in the C_i matri-

ces. For instance, one very special class of systems that can also be described by the general formalism of Eq. (2) is that of *temporal*, or time-varying networks. A temporal network is a system where the connections over a fixed set of some N nodes can fluctuate over time. They are described by a sequence of corresponding adjacency matrices $\boldsymbol{A} = \{A^{[1]}, A^{[2]}, \ldots, A^{[T]}\}$, where $A^{[t]}$ is the adjacency matrix of the network at a 'time' (or a position in the order of observation) $t = 1, 2, \ldots, T$. A temporal network can therefore be viewed as a multiplex network $\mathcal{M}$ as in Eq. (2), which has T layers ($M = T$), and where each layer α describes the system at 'time' $\alpha = t$. Notice that, in this case, the precise order of the matrices in $\boldsymbol{A}$ is important, since $A^{[\alpha]}$ and $A^{[\alpha+1]}$ represent the links of the time-varying system at two consecutive time observations. $\boldsymbol{C}$ also has a special structure since, for each node i, we have $c_i^{[\alpha\beta]} = c \cdot \delta_{\alpha+1,\beta}$ where $\delta_{\alpha\beta}$ is the Kronecker delta. Because of their very peculiar constraints and features, as well as the wide extent of literature on the topic, temporal networks will not be thoroughly discussed in this thesis. Information on this topic can be found in a variety of already available material, such as Ref. [8, 36, 37].

Multiplex, multilayer, interconnected and interdependent networks. We observe that a system described by eqs. (1) or (2) is a particular case within a generic tensorial formalism that encodes all possible connections in a systems with many layers in a rank four tensor $\mathcal{T} = \{\tau_{ij}^{[\alpha\beta]}\}$ [12]. As in the previous cases, superscripts indicate layers while subscripts indicate nodes, and the entry $\tau_{ij}^{[\alpha\beta]}$ is a non-negative real number representing the weight of a link from node i at layer α to node j at layer β. $\mathcal{T}$ thus stands for a general *multilayer* system, and can represent *interdependent* and *interconnected* networks that allow for inter-layer links between different nodes. In this notation multiplex systems are given as $\mathcal{M} \equiv \mathcal{T}$ supplemented by a requirement that, for $\alpha \neq \beta$, only entries $\tau_{ij}^{[\alpha\beta]}$ with $i = j$ can be different from zero. Attempts to provide a rigid classification of multilayer systems into multiplex, interconnected and interdependent networks are present in the literature [10]. Interconnected and interdependent networks have so far found a limited number of practical applications. We note here the case of the internet communication network and the network of power stations, where the effect of their respective breakdowns are multiplied by their interconnectedness [24], or the case of epidemic spreading in populations of males and females characterised by both homosexual and heterosexual interactions [38]. Conversely, the vast majority of multilayer networks considered in empirical and theoretical analysis follow into the category of multiplex networks, as described in vectorial formalism in eq. (1) or eq. (2). In this thesis we focus on multiplex networks, to which we also refer generically as multilayer networks. As network theory is not a tool which should belong to mathematicians and theoretical physicists only, but it is of great interest for people working on a variety of different and more applied fields, we will focus here on the vectorial formalism for

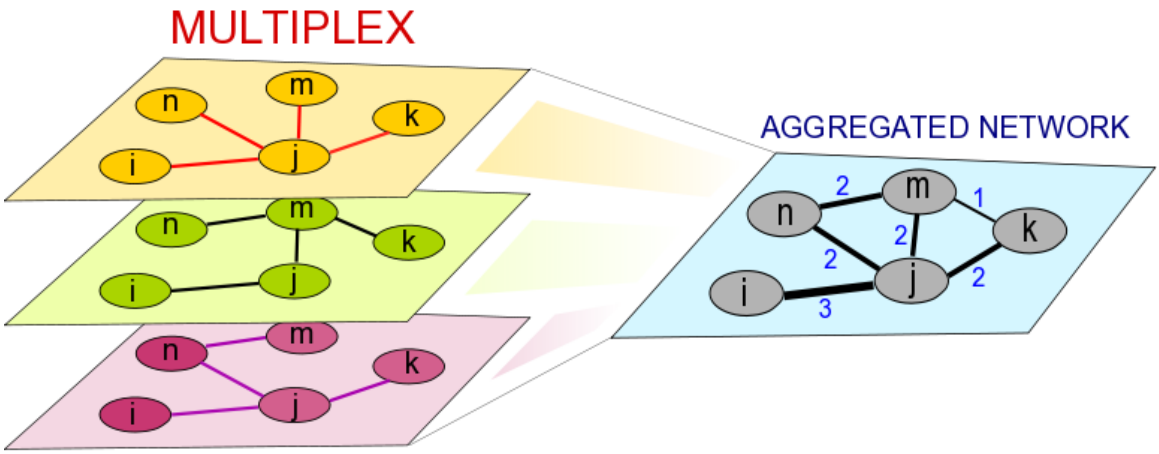


Figure 0.2: Schematic representation of a multiplex network with $N = 5$ nodes and $M = 3$ layers, corresponding to different types of relations (left). From a multiplex network, it is possible to obtain simplified structures, such as the weighted overlapping network $mathcalO$, by collapsing all layers and neglecting the different nature of the interactions (right).

multiplex networks, which we believe represent the simplest, yet the most effective way to describe multilayer networks.

Aggregated networks. We finally note that it is always possible to compress the information stored in a multiplex networks into simpler structures [15]. We first define the aggregated topological adjacency matrix $\mathcal{A} = \{a_{ij}\}$, where

$$a_{ij} = \begin{cases} 1 & \text{if } \exists\, \alpha : a_{ij}^{[\alpha]} = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Such a matrix describes a single-layer binary network where two nodes are connected if they share a link on at least one layer. It is known as the topological aggregated matrix and can be studied using the well-established set of measures defined for single-layer networks. In the topological aggregated matrix not only the nature of links connecting two nodes, but also their number, is lost in the aggregation process. This second basic feature can be preserved by introducing the overlapping network $\mathcal{O} = \{o_{ij}\}$, a single-layer network whose entries $o_{ij} = \sum_{\alpha=1}^{M} a_{ij}^{[\alpha]}$ accounts for the total number of edges connecting i and j across the different layers. In the next chapters of the thesis we will show that aggregated networks, either $\mathcal{A}$ or $\mathcal{O}$, often provide a non adequate description of networks with multiple types of interactions, falling short in capturing crucial information which can only be obtained by keeping track of the multiplexity characterising these systems.

Chapter 1

Node properties

When analysing a network, nodes are historically at the heart of our investigations. In social networks we are interested in discovering which individuals are the most central and influential. In brain networks we investigate which regions are crucial to process and transmit information. In transportation networks we want to know which stations are the hubs and should be avoided on peak time because they are likely to saturate. In economic networks we want to know which countries are in a position of advantage by looking at their import and export patterns of goods. In ecological networks we want to know which species are key to determine the stability of an ecosystem and its food chain. Node properties can be based on local information, such as the degree, or more complicated global information, for instance to determine its centrality. A striking feature of multiplex networks is that, differently from the traditional single-layer approach, where a generic property of a node i is described by a scalar variable, e.g. ξ_i, node properties are naturally described in vectorial terms, $\boldsymbol{\xi}_i = \{\xi_i^{[1]}, \ldots, \xi_i^{[M]}\}$. A basic starting point to the multiplex analysis of a system is to perform a layer-by-layer exploration of a given node property. Unfortunately, in the case of systems with many layers of interactions the information provided by this task is often difficult to read. As a consequence, one of the challenges is to describe the multiplexity characterising the units of the system by mean of synthetic descriptors which effectively combine information from the different layers. In this chapter we introduce and review a wide set of basic local and global properties for the nodes of a network in multilayer system.

1.1 Node activity

A crucial empirical evidence is that in many multiplex networks not all nodes have connections at all layers. As a consequence, a node i is defined as active on a layer α if it is connected to at least another node at that layer. The *activity*-pattern of each node can be compactly stored into the node-activity vector

$$\boldsymbol{b}_i = \{b_i^{[1]}, \ldots, b_i^{[M]}\}, \tag{1.1}$$

where $b_i^{[\alpha]} = 1 - \delta_{0,k_i^{[\alpha]}}$, i.e. $b_i^{[\alpha]} = 1$ if node i is active on layer α, and $b_i^{[\alpha]} = 0$ otherwise. The total activity $B_i = \sum_{\alpha=1}^{M} b_i^{[\alpha]}$ represents the number of layers in which node i is active, with $0 \leq B_i \leq M$ [22]. We indicate as $N^{[\alpha]}$ the number of nodes active, i.e. with at least one link, on layer α.

We perform an investigation of this basic feature of multiplex networks by introducing a number of different multiplex datasets, namely

- the six continental airline transportation systems, where airports are the nodes, routes between two airports are links, and different airlines are the layers (data available from OpenFlight at [39]);
- the American Physical Society (APS) co-authorship network, where scientists are the nodes, links indicate the existence of co-authored papers between two scientists, and different layers represent publications in ten different fields of physics, identified by their PACS code (data available at [40]);
- the Internet Movie Database (IDMb) co-starring network, where actors are the nodes, links indicate that two actors participated in the same movie, and different layers represent different movie genres (data available at [41]).

An important point is that related datasets were among the first one to be analysed in network science as single-layer networks, and the aim is to now gain some more insights on their structural patterns by rediscovering them in a multiplex context. In particular, pioneering work on the single-layer structure of the airport transport system can be found in Ref. [7, 42]. Conversely, early work on the structure of co-autorship and co-starring networks can be found respectively in Ref. [43] and in Ref. [5].

The basic properties of the six continental airline transportation systems and the two social networks are reported in Table 1-A. A detailed presentation of the layers of the APS and the IMDb multiplex networks is instead shown in Table 1-B and in Table 1-C respectively.

Network	N	M	$\langle N^{[\alpha]} \rangle$
Airlines - Africa	235	84	9.8
Airlines - Asia	792	213	24.4
Airlines - Europe	593	175	21.8
Airlines - North America	1020	143	24.9
Airlines - Oceania	261	37	14.1
Airlines - South America	296	58	15.1
APS	170385	10	43188
IMDb	2158300	28	229330

Table 1-A: Number of nodes N, number of layers M and average number of active nodes $\langle N^{[\alpha]} \rangle = \frac{1}{M} \sum_\alpha N^{[\alpha]}$ of the six continental airline transportation systems from OpenFlight [22].

Layer	**Field**	$N^{[\alpha]}$	$K^{[\alpha]}$	$\langle k^{[\alpha]} \rangle$
0	General	53170	1268045	47.7
1	Particles	37861	4865557	257.0
2	Nuclear	32792	1747892	106.6
3	Atomic	33649	189674	11.27
4	Classical	40269	222328	11.04
5	Gases and Plasmas	14237	179786	25.3
6	Condensed Matter I	63560	611765	19.3
7	Condensed Matter II	79416	631159	15.9
8	Interdisciplinary	45385	509058	22.4
9	Astronomy	31540	2467703	156.5

Table 1-B: The APS multiplex collaboration network consists of ten layers, one for each field of physics. For each layer α we report the number of active nodes $N^{[\alpha]}$, the number of edges $K^{[\alpha]} = \frac{1}{2} \sum_\alpha k^{[\alpha]}$ and the average degree $\langle k^{[\alpha]} \rangle$ [22].

As shown in Fig. 1.1, it has been found that these real-world multiplex networks are characterised by power-law distributions of node activity [22]. This means that the typical number of layers in which a node is active is subject to unbounded fluctuations, and that as a consequence its average value is not representative of the whole distribution.

1.2 Overlapping degree and participation coefficient

Another basic node property is the node *degree*, which accounts for the total number of connections of a node. As for the activity, also the degree of a node in a multiplex

Layer	**Genre**	$N^{[\alpha]}$	$K^{[\alpha]}$	$\langle k^{[\alpha]} \rangle$
1	Action	330333	11800436	71.4
2	Adult	66756	1691208	50.7
3	Adventure	210293	7390148	70.3
4	Animation	55376	1120523	40.5
5	Biography	128552	4272197	66.5
6	Comedy	810693	30118775	74.3
7	Crime	297554	10051325	67.6
8	Documentary	313019	6850670	43.8
9	Drama	1091789	43352371	79.4
10	Family	198301	5432262	54.8
11	Fantasy	176080	5096872	57.9
12	Film-Noir	7035	399548	113.6
13	Game-Show	15222	282942	37.2
14	History	124803	4137162	66.3
15	Horror	263290	5428250	41.2
16	Musical	121471	4118346	67.8
17	Music	165110	4977063	60.3
18	Mystery	168898	4226618	50.0
19	News	21530	406166	37.7
20	Reality-TV	29112	465244	32.0
21	Romance	364042	13325687	73.2
22	Sci-Fi	164468	4147689	50.4
23	Short	644430	5117780	15.9
24	Sport	101006	3643330	72.1
25	Talk-Show	19700	516943	52.5
26	Thriller	356776	10757551	60.3
27	War	118960	3967033	66.7
28	Western	56638	2101057	74.2

Table 1-C: Basic features on each of 28 layers of the IMDb multiplex network. In this case, each layer corresponds to a movie genre [22].

network is in general a vector

$$\boldsymbol{k}_i = \{k_i^{[1]}, \ldots, k_i^{[M]}\}, \tag{1.2}$$

where $k_i^{[\alpha]} = \sum_{i \neq j} a_{ij}^{[\alpha]}$ is the number of links incident in i at layer α.

We continue our exploration of multiplexity in real-world systems by introducing a fourth multilayer dataset, one of the first one considered in the literature, the Noordin Top Terrorist Network, first presented in Ref. [44] and deeply analysed in Ref. [15]. This dataset includes information about trust (T), operational (O), communication (C) ties and business (B) relations among a group of 78 terrorists from Indonesia active in recent years. Moreover, information for some of the layers can be split into a deeper level. This

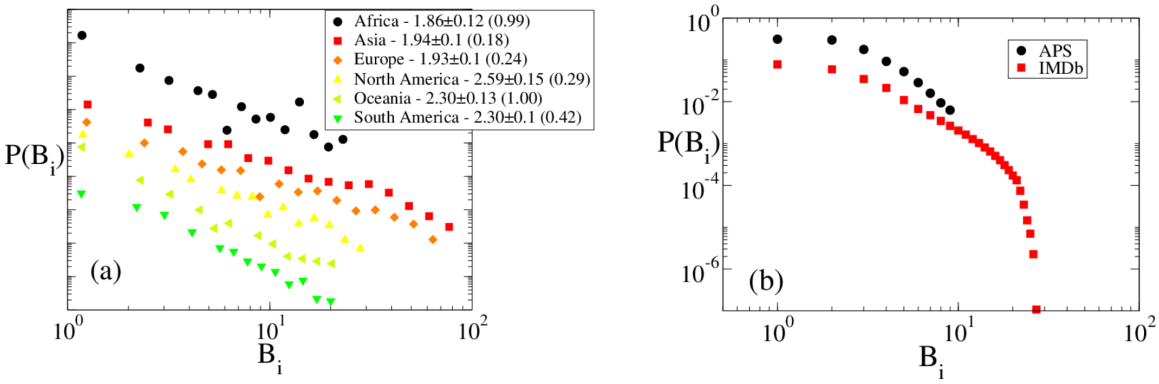


Figure 1.1: Distributions of node-activity for (a) the six multiplex networks of continental airlines and for (b) APS and IMDb. In all airline networks $P(B_i)$ can be fitted by power-laws with exponents ranging from 1.8 to 2.3. Exponents and corresponding p-values are reported in the legend in parenthesis [22].

is the case of the trust and operational networks which are composed by four sub-layers each, making them multiplexes inside a multiplex. Layer T is obtained as superposition of classmates, friendship, kinship and soul-mates ties, while layer O can be split into logistic, meetings, operations and training sub-layers. As a first approach we represent this system as a multiplex network with $M = 4$ layers, namely T, O, C and B. We exploit the additional richness of the data set to assign a weight to the links connecting nodes in layers T and O, while we leave the analysis of multiplexes of multiplexes for future work. In particular, we associated an integer number $w_{ij}^{[\mathrm{T}]}$ with $1 \leq w_{ij}^{[T]} \leq 4$ to every edge in the trust layer, based on how many times the connection appears in the four corresponding sub-layers. Analogously, an integer weight $w_{ij}^{[\mathrm{O}]}$ with $1 \leq w_{ij}^{[\mathrm{O}]} \leq 4$ is associated to every edge in the operational layer. For most of the following analysis we will consider also T and O as unweighted layers. An analysis of the weights of the links on layers T and O is presented in the next chapter.

Summing up, the multiplex network of Top Noordin Terrorists has $N = 78$ nodes, $K = 623$ (where K is the number of links in the topologically aggregated binary networks), $O = 911$ (where O is the sum of the number of links of each layer) and $O^{\mathrm{w}} = 1014$ (where O^{w} is the sum of the weights of the links of each layer). Table 1-D reports more details about the size of each layer and sub-layer, and of the corresponding aggregated adjacency matrices. We notice that some individuals are not involved in all the four layers, meaning that their activity with respect to a particular kind of social relationship has not been registered or was unknown at the time the data set was compiled. Consequently, some of the replicas of such nodes will be isolated nodes on one or more of the four layers. It is evident that while the trust, communication and operational layers share approxi-

LAYER	CODE	N	K	S	O	O^{w}
MULTIPLEX	M	78	623	/	911	1014
TRUST	T	70	259	293	/	/
Classmates	Tc	39	175	/	/	/
Friendship	Tf	61	91	/	/	/
Kinship	Tk	24	16	/	/	/
Soulmates	Ts	9	11	/	/	/
OPERATIONAL	O	68	437	506	/	/
Logistic	Ol	16	29	/	/	/
Meetings	Om	26	63	/	/	/
Operations	Oo	39	267	/	/	/
Training	Ot	38	147	/	/	/
COMMUNICATION	C	74	200	200	/	/
BUSINESS	B	13	15	15	/	/

Table 1-D: The Top Noordin Terrorist Network includes data about trust (T), operations (O), communication (C) and business (B) among 78 terrorists active in recent years in Indonesia. Trust and operational networks are characterized by a deeper internal structure, and they can be divided into four sub-layers each. For the multiplex network (M), and each layer and sub-layer we show the total number of active nodes N, and the number of edges expressed as non-overlapping links K (number of links in the aggregated topological network), overlapping links O (sum of the number of links of the four layers, where each layer is considered as unweighted) and weighted overlapping links O^{w} (sum of the number of links of the four sublayers for trust, the four sublayers for operational, together with communication and business). For each layer α we also report the total strength $S = \frac{1}{2} = \sum_i s_i$ [15].

mately 90% of the nodes, the business layer has only 13 active nodes. Consequently, in the following we will usually consider only trust, communication and operational relationships. For this three-layer multiplex network we have $N = 78$, $K = 620$ (total number of links in the aggregated network), $O = 896$ (total number of links obtained by summing the layers of the multiplex networks) and $O^{\mathrm{w}} = 999$ (total weight obtained by summing the weights of all links). A schematic representation of this multiplex network is reported in Fig. 1.2. We introduce the total number of connections of node i, usually known as total or *overlapping degree* [15], as

$$o_i = \sum_{\alpha=1}^{M} k_i^{[\alpha]}. \tag{1.3}$$

The node color-code indicates the layers in which nodes are involved, while the size of

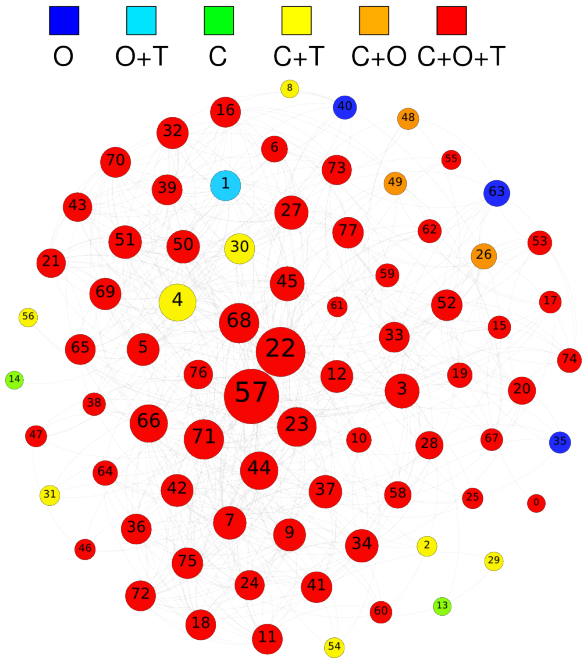


Figure 1.2: A flattened representation of the three-layer multiplex obtained by considering only trust (T), communication (C) and operations ties (O). For each node i we indicate with a color-code the layers in which i is actively involved (i.e., the layers α for which $k_i^{[\alpha]} > 0$). The size of a node is proportional to its overlapping degree o_i: node 57 is the node with the largest overlapping degree [15].

each node is proportional to its overlapping degree o_i. Notice that in this case most of the nodes participate to all the three layers, while just a few of them are present in only one or two layers.

We can now evaluate how the degree of a node is distributed across different layers. It is in fact possible that nodes which are hubs in one layer have only few connections, or are even isolated, in another layer. Or, alternatively, nodes which are hubs in one layer are also hubs in the other layers. We have therefore computed the aggregated topological degree k_i and the degree of the nodes in each layer $k_i^{[\alpha]}$, with $\alpha \in \{\mathrm{T}, \mathrm{O}, \mathrm{C}\}$, ranking the nodes according to their aggregated topological degree. In Fig. 1.3(a) we compare with a color-code plot the values of k_i, the degree on the topological aggregated networks, with the values $k_i^{[\alpha]}$ of the node degree at each layer α. By visual inspection, the four degree sequences appear weakly correlated, with nodes which are hubs in one level often having only few connections in another layer. In Fig. 1.3(b) we report the results obtained by ranking the nodes according to o_i, their total number of connections in the weighted overlapping network $\mathcal{O}$. Also in this case we observe weak correlations

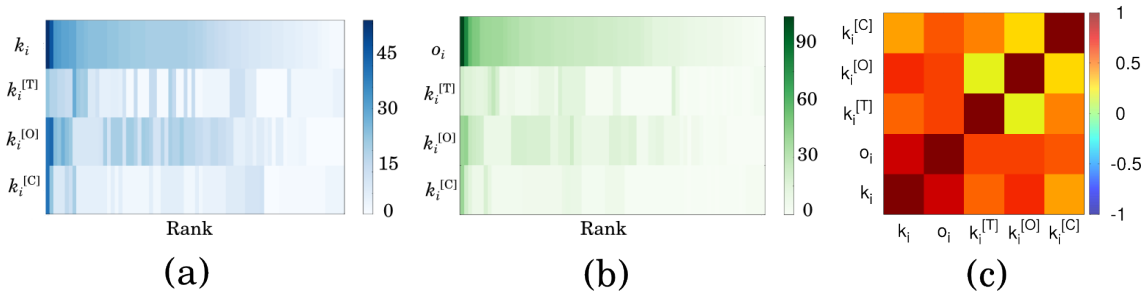


Figure 1.3: The top row of the panel shows with a color-code the degree k_i of each node, from the largest (darkest, leftmost) to the smallest (brightest, rightmost). Keeping fixed the ranking induced by k_i, in the other three rows we report respectively the degree in the trust layer $k_i^{[\mathrm{T}]}$, operational layer $k_i^{[\mathrm{O}]}$ and communication layer $k_i^{[\mathrm{C}]}$. (b) Same as panel (a) but in the first row nodes are ranked according to their overlapping degree o_i. (c) The heat map represents the Kendall τ correlation coefficient between k_i, o_i, $k_i^{[\mathrm{T}]}$, $k_i^{[\mathrm{O}]}$ and $k_i^{[\mathrm{C}]}$. Notice that the degree of a node in the operational layer O is poorly correlated with its degree in the communication and trust layers (bright yellow regions in the heat map) [15].

between the four degree sequences. To better quantify such correlations, we computed the Kendall rank correlation coefficient, τ_k, which measures the similarity of two ranked sequences of data X and Y. The correlation coefficient τ_k is a non-parametric measure of statistically dependence between two rankings, since it does not make any assumption about the distributions of X and Y, and takes values in $[-1, 1]$. We get $\tau_k(X, Y) = 1$ if the two rankings are identical, $\tau_k(X, Y) = -1$ if one ranking is exactly the reverse of the other and finally $\tau_k(X, Y) = 0$ if X and Y are independent. In Fig. 1.3(c) we report as a heat map the values of τ_k obtained for the rankings of each pair of variables. Notice that the aggregated degrees k_i and o_i are usually weakly correlated with the degree of node i on each single layer. The highest correlation is indeed found between the degree of the aggregated topological network k_i and the overlapping degree o_i.

Given a generic vectorial property $\boldsymbol{\xi}_i$, it is important to be able to compress the information into meaningful scalar descriptors, especially for systems composed of a large number of layers. A typical way to approach this problem is to consider the first and the second moment of the vector $\boldsymbol{\xi}$, accounting for its mean value $\mu(\boldsymbol{\xi})$ (or, analogously, the sum of its components) and its variance $\sigma^2(\boldsymbol{\xi})$, or related quantities. In the particular case of the degree, two synthetic descriptors have been successfully introduced in the literature. The first one is the previously introduced total number overlapping degree $o_i = \sum_{\alpha=1}^{M} k_i^{[\alpha]}$, which accounts for the total number of connections of node i disregarding their specific nature [15]. Conversely, the heterogeneity of the

number of neighbours of node i across the layers can be measured through the multiplex *participation coefficient* [15]

$$P_i = \frac{M}{M-1}\left[1 - \sum_{\alpha=1}^{M}\left(\frac{k_i^{[\alpha]}}{o_i}\right)^2\right], \tag{1.4}$$

where $P_i = 1$ when the links incident on node i are equally distributed across the layers, and $P_i = 0$ when a node is only active on one layer. The definition of the multiplex participation coefficient is in the same spirit of that of participation coefficient introduced in Refs. [45, 46] to quantify the participation of a node to the different communities of a network. We note that similar information about the heterogeneity of the distribution of a node's connections across layers is provided by the Shannon entropy of the degree vector [15]

$$H_i = -\sum_{\alpha=1}^{M}\frac{k_i^{[\alpha]}}{o_i}\ln\left(\frac{k_i^{[\alpha]}}{o_i}\right). \tag{1.5}$$

In Fig. 1.4(a) we plot the distribution of P_i for the multi-layer network of Indonesian terrorists under study. Although the average participation coefficient of the multiplex is equal to $P = 0.72$, we observe a quite broad distribution of P_i in the range $[0, 1]$. This variance suggests the existence in the network of various levels of node participation to each of the three layers. Since the overlapping degree of a node represents its overall importance in terms of number of incident edges, while the multiplex participation coefficient gives information about the distribution of incident edges across the layers, we propose to classify the nodes of a multiplex by looking, at the same time, at their multiplex participation coefficient and at their overlapping degree. With respect to the multiplex participation coefficient, we identify three classes. We call *focused* those nodes for which $0 \leq P_i \leq 1/3$, *mixed* the nodes having $1/3 < P_i \leq 2/3$ and *truly multiplex* (or even simply *multiplex*) the nodes for which $P_i > 2/3$. Instead of the overlapping degree we consider the associated Z-score, which allows to compare multiplex networks of different size:

$$z(o_i) = \frac{o_i - \langle o\rangle}{\sigma_o} \tag{1.6}$$

where $\langle o\rangle$ is equal to the average overlapping degree of the nodes of the system, and σ_o is the corresponding standard deviation. With respect to the Z-score of their overlapping degree, we distinguish *hubs*, for which $z(o_i) \geq 2$, from regular nodes, for which $z(o_i) < 2$. Consequently, by considering the multiplex participation coefficient P_i of a node and its total overlapping degree o_i we can define six classes of nodes, as depicted in Fig. 1.4(b), where we represent each node as a point in the $(P_i, z(o_i))$ plane.

Notice that the distribution of $z(o_i)$ is asymmetric and unbalanced towards positive

values, and this is a sign of the heterogeneity of the total overlapping degree. Moreover, there is a quite large heterogeneity in the values of P_i for a fixed value of $z(o_i)$. Let us focus for instance on two specific nodes, namely node 16 and 34. These two nodes have the same overlapping degree, namely $o_{16} = o_{34} = 25$, corresponding to $z(o_{16}) = z(o_{34}) = 0.12$, but very different participation coefficient across layers T, O and C, respectively $P_{16} = 0.915$ and $P_{34} = 0.23$. Consequently, even if the overall number of edges of node 16 and node 34 is the same (which would make these two nodes indistinguishable in the aggregated overlapping network), they play radically different roles, as becomes evident by looking at their ego networks, reported in Fig. 1.4(c). In fact, while node 34 is highly focused on the operational layer (blue edges), with only one edge in the trust layer (green edge) and one edge in the communication layer (red edge), node 16 is instead involved in all the three layers, with a comparable number of edges in each of them. This implies that the removal of node 34 would primarily affect just the operational layer, while the absence of node 16 could cause major disruptions in the trust, operational and communication networks. Similar results are obtained by considering the Z-score of the degree k_i of node i in the aggregated topological network (figure not shown).

To conclude, the pair of variables (P_i, o_i) can be used to classify nodes via the so-called multiplex cartography [15], efficiently distinguishing multiplex hubs (high o_i and high P_i), focused hubs (high o_i and low P_i), multiplex leaves (low o_i and high P_i) and focused leaves (low o_i and low P_i). The overlapping degree and the participation coefficient can be easily generalised to the case of networks with weighted layers, where the degree of a node is replaced by its *strength* $\boldsymbol{s}_i = \{s_i^{[1]}, \ldots, s_i^{[M]}\}$, with $s_i^{[\alpha]} = \sum_{j \neq i} w_{ij}^{[\alpha]}$.

1.3 Clustering

One of the most remarkable characteristic of complex real-world single-layer networks, especially acquaintance and collaboration networks, is the tendency of nodes to form triangles, i.e. simple cycles involving three nodes. This widely observed tendency is concisely expressed by the popular saying *"the friend of your friend is my friend"* and is usually quantified through the so-called node *clustering coefficient* [5]. The clustering coefficient of node i is defined as :

$$C_i = \frac{\sum_{j \neq i, m \neq i} a_{ij} a_{jm} a_{mi}}{\sum_{j \neq i, m \neq i} a_{ij} a_{mi}} = \frac{\sum_{j \neq i, m \neq i} a_{ij} a_{jm} a_{mi}}{k_i(k_i - 1)}. \tag{1.7}$$

and quantifies how likely it is that two neighbors of node i are connected to each other. In fact, Eq. (1.7) measures the fraction of triads centered in i that close into triangles. By definition C_i takes values in the interval $[0, 1]$. Averaging this quantity over all the

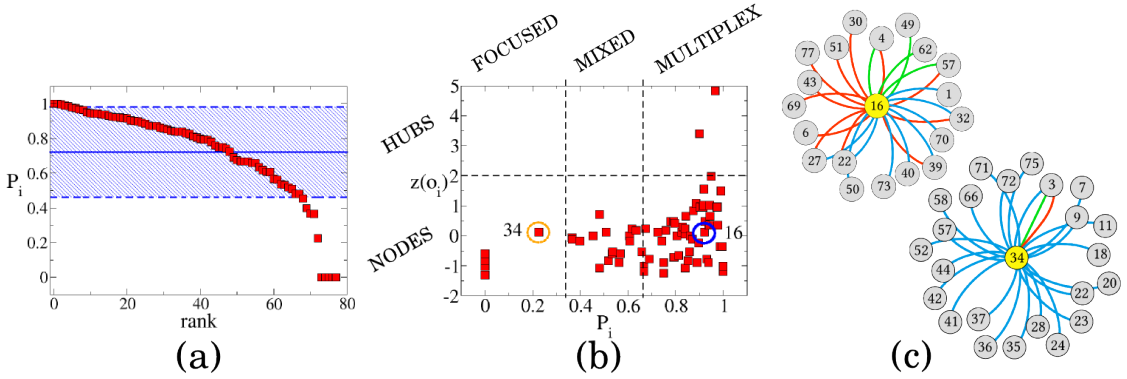


Figure 1.4: (a) Rank distribution of the participation coefficient P_i for the multi-layer network of Top Noordin Indonesian Terrorists. $M = 3$ layers were considered, namely trust, operational and communication. The average value $P = 0.72$ is shown as a horizontal blue line, while the shaded band indicates the standard deviation. (b) A cartography of the roles of the nodes in a multi-layer network can be obtained by plotting, for each node i, the multiplex participation coefficient P_i versus the Z-score of the total overlapping degree $z(o_i)$. Even if two nodes have exactly the same value of $z(o_i)$ (like node 16 and node 34, indicated by the orange and blue circles, respectively), they can have pretty different roles, according to the value of the multiplex participation coefficient. (c) The ego networks of node 16 and 34, in which edges are colored according to the layer to which they belong, respectively green (trust), blue (operational) and red (communication). It is evident that the connectivity pattern of node 16, whose links are homogeneously distributed across the three layers, is "more *multiplex*" than that of node 34, which is instead *focused* on the operational layer [15].

nodes in a network, one gets the network clustering coefficient:

$$C = \frac{1}{N} \sum_i C_i. \tag{1.8}$$

A similar —although not identical— measure of local cohesion [47], which is commonly used in the social sciences, is the network *transitivity* [48]:

$$T = \frac{3 \times \text{No. of triangles in the graph}}{\text{No. of triads in the graph}}. \tag{1.9}$$

This is defined as the proportion of triads, i.e. connected triples of nodes, which close into triangles.

Since each layer of a multiplex can be seen as a single-layer network, the definitions of network clustering coefficient and network transitivity can be used to characterize the

abundance of triangles on each layer. In general, different layers may show similar or dissimilar patterns of clustering. In Table 1-E we report the average clustering coefficient and the transitivity for each layer of the terrorist network, and for its topological aggregate.

Layer	C	T
T	0.38	0.53
O	0.67	0.62
C	0.45	0.27
$\mathcal{A}$	0.66	0.56

Table 1-E: The average clustering coefficient C and the transitivity T for layers T, O, C and for the aggregated topological network $\mathcal{A}$ for the Noordin dataset [15].

Notice that each layer has quite peculiar values of clustering and transitivity, which are in turn different from those measured on the aggregated topological network. In particular, the highest values of clustering and transitivity are observed in the Operations layer, probably due to the fact that terrorist missions usually involve more than two people at the same time. In Fig. 1.5(a) we focus on the node clustering coefficient, we rank the nodes of the multiplex according to the value of C_i for the aggregated topological network and we compare this value with the clustering coefficient calculated on each layer $C_i^{[\alpha]}$. As shown, many nodes display quite different values of the clustering coefficient across the layers. We have computed the Kendall correlation coefficient τ_k between each pair of layers and between each layer and the topological aggregate. The results are shown in Fig. 1.5(b), as a heat map. Notice that at the best the sequences of clustering coefficient are weakly correlated, when not uncorrelated or even anti-correlated. In particular, the ranking of clustering coefficient for the Operations layer is anti-correlated with that of the other three layers and of the topological aggregated network.

However, comparing the sequences of C_i for each layer tells us very little about the interplay between the several levels of the system in terms of clustering. In particular, it is interesting to study to which extent the multiplexity affects the formation of triangles, i.e. how the presence of different layers can give rise to triangles which were impossible to close at the level of single layers. For this reason we need to extend the notion of triangle to take into account the richness added by the presence of more than one layer. We define a *2-triangle* a triangle which is formed by an edge belonging to one layer and two edges belonging to a second layer. Similarly, we call a *3-triangle* a triangle which is composed by three edges all lying in different layers. In order to quantify the added value provided by the multiplex structure in terms of clustering, we define two parameters of *clustering interdependence* I_1 and I_2. I_1 is the ratio between the number of triangles in

the multiplex which can be obtained only as 2-triangles, and the number of triangles in the aggregated system. I_2 is the ratio between the number of triangles in the multiplex which can be obtained only as 3-triangles and the number of triangles in the aggregated system. Then, $I = I_1 + I_2$ is the total fraction of triangles of the aggregated topological network which can not be found entirely in one of the layers. For the multi-layer network of terrorists we obtain $I_1 = 0.31$ and I_2 of the order of 10^{-3}, which indicates that almost no triangle is formed exclusively by the interplay of three different layers. This result is the first indicator of the presence of non-trivial patterns in clustering and triadic closure in multi-layer systems.

In this section we also aim at generalizing the notion of clustering coefficient to multi-layer networks. Recalling the definition of 2-triangle and 3-triangle, we define a *1-triad* centered at node i, for instance $j - i - k$, a triad in which both edge $j - i$ and edge $i - k$ are on the same layer. We also define a *2-triad* as a triad whose two links belong to two different layers of the systems. We are now ready to give two definitions of clustering coefficient for multiplex networks. The first coefficient $C_{i,1}$ is defined, for each node i, as the ratio between the number of 2-triangles with a vertex in i and the number of 1-triads centered in i. We can express this clustering coefficient in terms of the multi-layer adjacency matrix as:

$$C_{i,1} = \frac{\sum_\alpha \sum_{\alpha' \neq \alpha} \sum_{j \neq i, m \neq i} (a_{ij}^{[\alpha]} a_{jm}^{[\alpha']} a_{mi}^{[\alpha]})}{(M-1) \sum_\alpha \sum_{j \neq i, m \neq i} (a_{ij}^{[\alpha]} a_{mi}^{[\alpha]})} = \\ = \frac{\sum_\alpha \sum_{\alpha' \neq \alpha} \sum_{j \neq i, m \neq i} (a_{ij}^{[\alpha]} a_{jm}^{[\alpha']} a_{mi}^{[\alpha]})}{(M-1) \sum_\alpha k_i^{[\alpha]} (k_i^{[\alpha]} - 1)} \tag{1.10}$$

Since each 1-triad can theoretically be closed as a 2-triangle on each of the M layers of the multiplex excluding the layer to which its edges belong, in order to have a normalised coefficient we have to divide the term by $M - 1$. In addition to this, we define a second clustering coefficient for multiplex networks as the ratio between the number of 3-triangles with node i as a vertex, and the number of 2-triads centered in i. In terms of adjacency matrices, we have:

$$C_{i,2} = \frac{\sum_\alpha \sum_{\alpha' \neq \alpha} \sum_{\alpha'' \neq \alpha, \alpha'} \sum_{j \neq i, m \neq i} (a_{ij}^{[\alpha]} a_{jm}^{[\alpha'']} a_{mi}^{[\alpha']})}{(M-2) \sum_\alpha \sum_{\alpha' \neq \alpha} \sum_{j \neq i, m \neq i} (a_{ij}^{[\alpha]} a_{mi}^{[\alpha']})}. \tag{1.11}$$

where a normalisation coefficient $M-2$ has been added. While $C_{i,1}$ is a suitable definition for multiplexes with $M \geq 2$, $C_{i,2}$ can only be defined for systems composed of at least three layers. Averaging over all the nodes of the system, we obtain the network clustering coefficients C_1 and C_2. Similar definitions can also be provided in terms of the tensorial

formalism for multilayer networks, as in Ref. [49].

In Fig. 1.5(c) we rank the nodes of the terrorist network according to their value of C_i for the aggregated system, and compare this sequence of values with the ones obtained with the two measures of multiplex clustering, $C_{i,1}$ and $C_{i,2}$. As shown in the figure, $C_{i,1}$ and $C_{i,2}$ capture different effects of multi-layer clustering. This fact is confirmed by the heat map reported in Fig. 1.5(d), which shows with a color-code the non-parametric correlations among $C_{i,1}$, $C_{i,2}$ and C_i. Notice that, in general, the correlation between C_i and both $C_{i,1}$ and $C_{i,2}$ is pretty small.

These results indicate that multiplex clustering provides information which are substantially different from those obtained by looking at the clustering of the aggregated network. In addition to this, $C_{i,1}$ and $C_{i,2}$ are poorly correlated, as is also evident from Fig. 1.6(a). In practice, for a given value of $C_{i,1}$, we have nodes with a wide range of values of $C_{i,2}$, and vice–versa. Consequently, it is necessary to use both clustering coefficients in order to properly quantify the abundance of triangles in multi-layer networks. In Fig. 1.6(b) and Fig. 1.6(c) we report the scatter-plots of $C_{i,1}$ and $C_{i,2}$ versus o_i. Multiplex clustering coefficients are genuine multiplex variable and appear to be not correlated with the degree of the nodes of the system. We also found that the clustering coefficient is not correlated with other measures of aggregated degree, such as k_i.

We can also generalize the definition of transitivity T to the case of multi-layer networks. Similarly to the case of the clustering coefficient we propose two measures of transitivity. We define T_1 as the ratio between the number of 2-triangles and $M-1$ times the number of 1-triads in the multi-layer network. Moreover, we introduce T_2 as the ratio between the number of 3-triangles and $M-2$ times the number of 2-triads in the system.

Notice that clustering interdependences I_1 and I_2, average multiplex clustering coefficients C_1 and C_2 and multiplex transitivities T_1 and T_2 are all global graph variables which give a different perspective on the multi-layer patterns of clustering and triadic closure with respect to the clustering coefficient and the transitivity computed for each layer of the network. We have computed all such quantities for the multi-layer network of the Indonesian terrorists and, as a term of comparison, we have constructed a configuration model for multiplex networks, which will be useful to prove the non-trivial organization of the network under study.

In analogy with the case of a single-layer network, for a multiplex with M layers, where each node is characterized by a degree vector $\boldsymbol{k}_i$, we call configuration model the set of multiplexes obtained from the original system by randomizing edges and keeping fixed

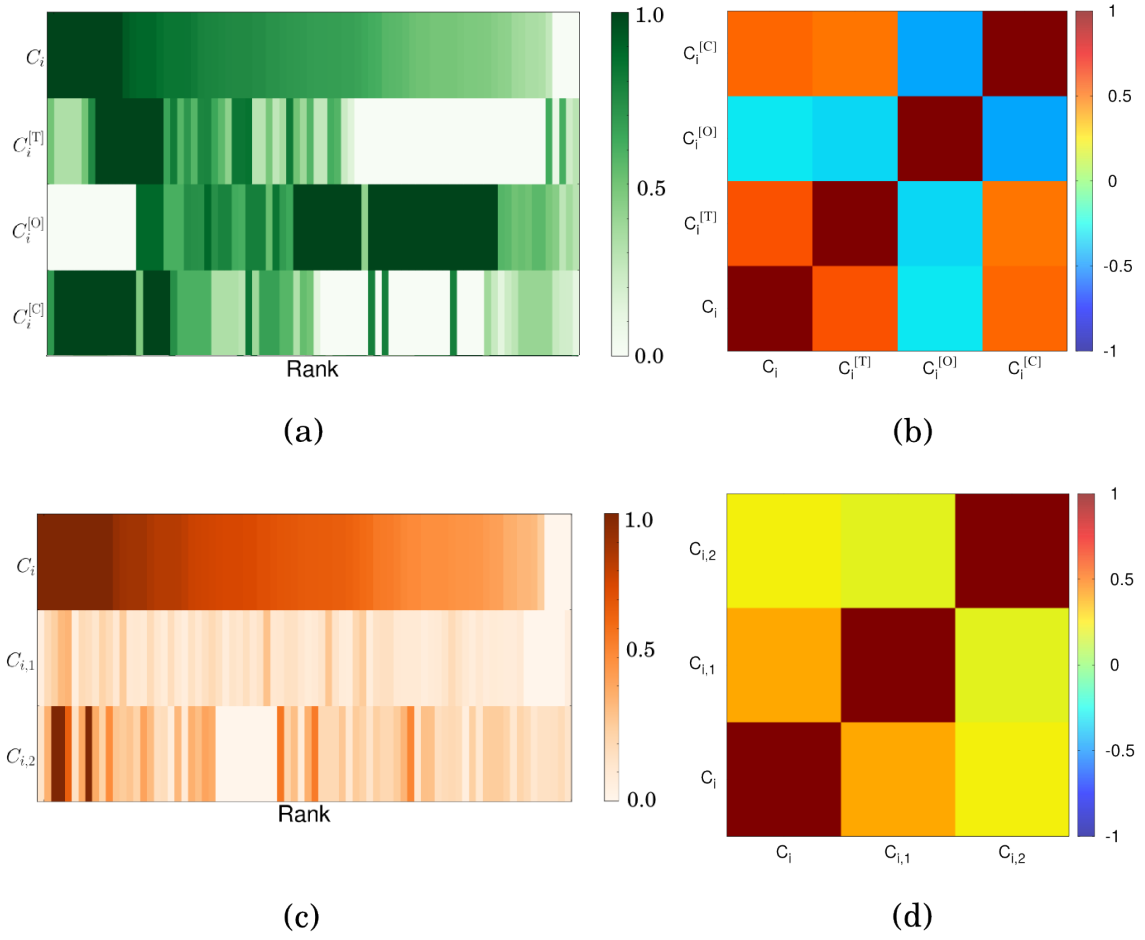


Figure 1.5: (a) The node clustering coefficient C_i of the aggregated topological network and of the three layers T, O, C, respectively denoted as $C_i^{[\mathrm{T}]}$, $C_i^{[\mathrm{O}]}$ and $C_i^{[\mathrm{C}]}$. The nodes are ranked according to their value of C_i on the aggregated topological network. (b) The heat map represents the correlation between the rankings of nodes according to their clustering coefficients on the three layers and on the topological aggregated network. Notice that $C_i^{[\alpha]}$ is weakly correlated with C_i for $\alpha \in \{T, O, C\}$, and that such correlation might also be negative, as in the case of $C_i^{[\mathrm{O}]}$. (c) Comparison among the clustering coefficient C_i of the aggregated topological network, and the multi-layer clustering coefficients $C_{1,i}$ and $C_{2,i}$. The nodes are ranked according to their value of C_i. (d) The heat map represents the correlation between the rankings of nodes according to C_i, $C_{1,i}$ and $C_{2,i}$ [15].

the sequence of degree vectors $\{\boldsymbol{k}_1, \boldsymbol{k}_2, \ldots, \boldsymbol{k}_N\}$, i.e. keeping fixed the degree sequence at each layer α. We can now compare the values of C and T, C_1 and C_2, T_1 and

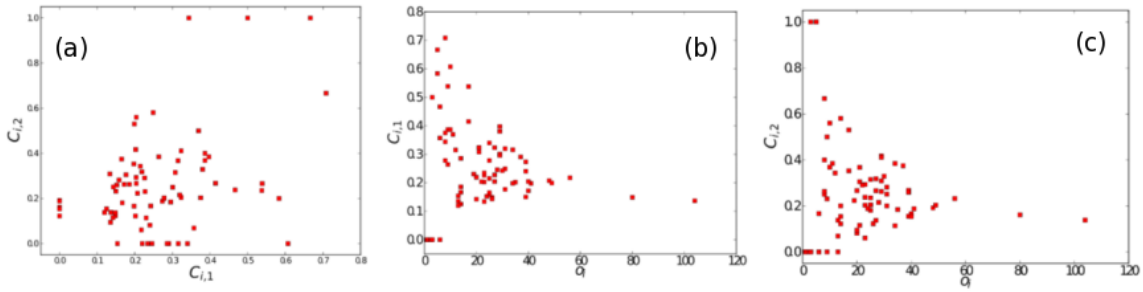


Figure 1.6: Scatter-plots of (a) $C_{1,i}$ versus $C_{2,i}$ (b) $C_{1,i}$ versus o_i and (c) $C_{2,i}$ versus o_i. The values of the Kendall's τ and of the Pearson's linear correlation coefficient r for any pair of measures are, respectively: $\tau(C_{i,1}, C_{i,2}) = 0.61$, $r(C_{i,1}, C_{i,2}) = 0.76$, $\tau(C_{i,1}, o_i) = -0.11$, $r(C_{i,1}, o_i) = -0.13$, $\tau(C_{i,2}, o_i) = 0.01$, $r(C_{i,2}, o_i) = 0.04$. It is worth noticing that both $C_{i,1}$ and $C_{2,i}$ are almost uncorrelated with the overlapping degree o_i, a fact that confirms their truly multiplex nature [15].

T_2, I_1 and I_2 obtained on real data with the average values found for the multi-layer configuration model. The comparison is shown in Table 1-F. As expected C and T computed on the aggregated topological network for real data are systematically higher than the ones obtained on randomized data, where edge correlations are washed out by the randomization. For the same reason, C_1, C_2, T_1 and T_2 are higher on real data. Conversely, we obtained higher values on randomized data for I_1 and I_2. This is not surprising, since the measures of clustering interdependence tell us about the fraction of triangles which can be exclusively found as multi-triangle in the system. We note that the randomisation process destroys comparatively more 1-triangles, which account for clustering inside each layer, than 2- or 3-triangles, which measure clustering across layers. In the randomised system, this leads to a decrease in the denominators of I_1 and I_2 which is much stronger than that of the corresponding numerators. We also finally notice that since the configuration model washes out inter-layer correlations, it is generally easier to find multi-triangles on a randomized multiplex network rather than on a real one where edges have a higher overlap. All these results demonstrate that, as previously shown for the overlap, also the clustering coefficient appears to be affected by the presence of non-trivial structural properties across the different layers of the multiplex network under study.

Variable	Real data	Randomized data
C	0.66	0.46
T	0.56	0.41
C_1	0.13	0.08
C_2	0.26	0.18
T_1	0.10	0.07
T_2	0.21	0.16
I_1	0.31	0.60
I_2	0.005	0.047

Table 1-F: Values of clustering C and transitivity T computed on the aggregated topological network, and values of the introduced measures for clustering in multi-layer networks, namely the multiplex clustering C_1 and C_2, the multiplex transitivity T_1 and T_2, and the clustering interdependence I_1 and I_2. For comparison we report also the results for a randomized system obtained through a multi-layer configuration model [15].

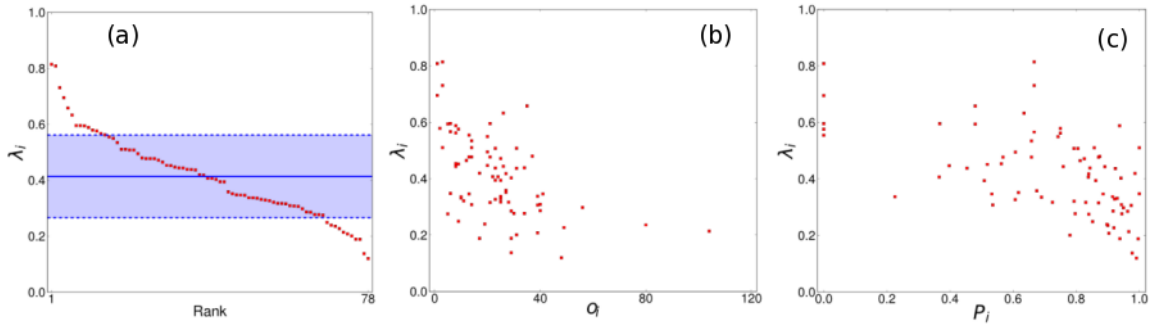


Figure 1.7: (a) Rank distribution of the node interdependence λ_i in the Indonesian Terrorist multiplex network. (b) Scatter-plot of the interdependence λ_i versus o_i and (c) versus P_i. The corresponding value of Kendall's τ and Pearson's r correlation coefficient are, respectively, $\tau(\lambda_i, o_i) = -0.41$, $r(\lambda_i, o_i) = -0.56$, $\tau(\lambda_i, P_i) = -0.41$, $r(\lambda_i, P_i) = -0.57$ [15].

1.4 Interdependence

Reachability is an important feature in networked systems. In single-layer networks it has to do with the existence and length of shortest paths connecting pairs of nodes. In multi-level systems, shortest paths may significantly differ between different layers, and each layer and the aggregated topological networks as well. To capture the multiplex contribution to the reachability of each unit of the network, the so-called node *interdependence* has been recently introduced in Ref. [50]. The interdependence λ_i of node i is

defined as:

$$\lambda_i = \frac{1}{N-1} \sum_{j \neq i} \frac{\psi_{ij}}{\sigma_{ij}}, \tag{1.12}$$

where σ_{ij} is the total number of shortest paths between node i and node j on the multiplex network, and ψ_{ij} is the number of shortest paths between node i and node j which make use of links in two or more than two layers. Hence, the node interdependence is equal to 1 when all shortest paths make use of edges laying at least on two layers, and equal to 0 when each of the shortest paths makes use of only one of the M layers of the system. Averaging λ_i over all nodes, we obtain the network interdependence $\lambda = (1/N) \sum_i \lambda_i$. In Fig. 1.7 we display the rank distribution of λ_i. The network has a large variety of node interdependencies: although most of the nodes have a value of λ_i in the range $[0.27, 0.56]$ around the average value $\lambda = 0.41$, there are also nodes with values as small as $\lambda_i = 0.1$, and two nodes with values larger than 0.8.

The interdependence is a genuine multiplex measure and, as shown in Fig. 1.7(b) provides information in terms of reachability which is slightly anti-correlated to measures of degree such as o_i. In fact, a node with a high overlapping degree quite likely will have a number of different possibilities to choose the first edge to go towards the other nodes, and in this way it will have a low value of λ_i. Conversely, a node with low degree will more likely have a high value of λ_i, being its shortest paths constrained to a limited selection of edges and layers from the first step. Moreover, λ_i appears to be slightly anti-correlated with P_i, as confirmed by the values of Kendall's and Pearson's correlation coefficients (see the caption of Fig. 1.7). We finally note that it is also possible to define a layer interdependence $\lambda^{[\alpha]}$, accounting for the number of shortest paths with at least one link on layer α [51].

1.5 Centrality

Another characteristic property of real-world networks is the presence of heterogeneity in the relative importance of nodes, as measured by different notions of node *centrality*. As for other basic properties, the centrality of nodes at different layers can be very different. For instance, let us consider the eigenvector centrality of node i in the multiplex network. As for the other node properties, it is a vector:

$$\boldsymbol{E}_i = \{E_i^{[1]}, \ldots, E_i^{[M]}\}. \tag{1.13}$$

where $E_i^{[\alpha]}$ is the eigenvector centrality of node i at layer α. We also indicate the eigenvector centrality on the aggregated topological and on the aggregated overlapping

network respectively as $E_i(\mathcal{A})$ and $E_i(\mathcal{O})$. In Fig. 1.8(a), 1.8(b) and 1.8(c) we compare the eigenvector centrality computed on each layer with that evaluated on the aggregated topological and overlapping networks. We notice only very weak correlations between the different centrality sequences. Such results are very similar to those obtained in Section 7.2 for the case of node degree, as a consequence of the fact that, at order zero, the eigenvector centrality reduces to the node degree. The Kendall correlation coefficients obtained for pairs of centralities are reported in Fig. 1.8(d) as a heat map.

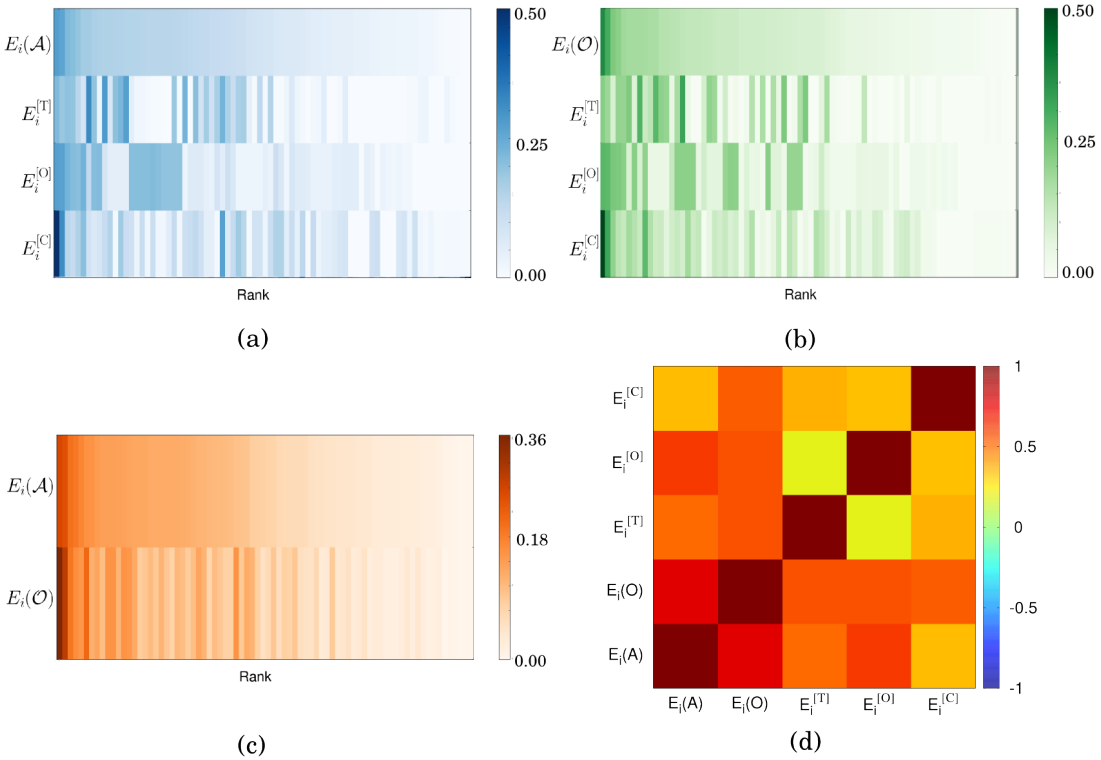


Figure 1.8: (a) Eigenvector centrality of the aggregated topological network $E_i(\mathcal{A})$, and of the trust $E_i^{[\mathrm{T}]}$, operational $E_i^{[\mathrm{O}]}$ and communication layer $E_i^{[\mathrm{C}]}$. The nodes are ranked according to their value of $E_i(\mathcal{A})$ on the aggregated topological network. (b) Similar to panel (a) but here nodes are ranked according to their eigenvector centrality computed on the aggregated overlapping network $E_i(\mathcal{O})$. (c) Comparison of the rankings of eigenvector centrality computed on the aggregated topological network and on the aggregated overlapping network, respectively $E_i(\mathcal{A})$ and $E_i(\mathcal{O})$. (d) The heat map shows the non-parametric correlation between the rankings induces by the different centralities [15].

For a large fraction of nodes, the rankings induced by the eigenvector centrality at

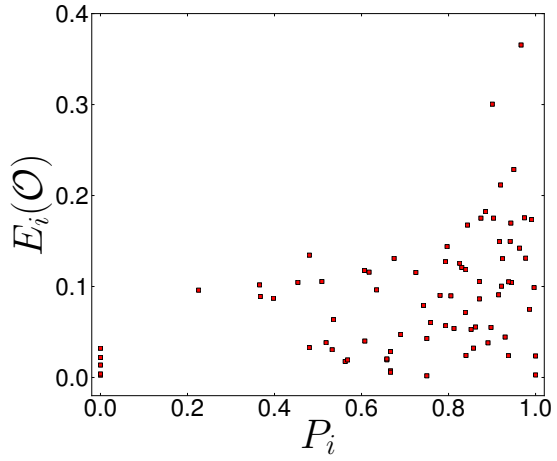


Figure 1.9: Scatter-plot of the eigenvector centrality on the aggregated overlapping network $E_i(\mathcal{O})$, and the participation coefficient P_i. Notice that there is indeed a slightly positive correlation between these metrics ($\tau(E_i(\mathcal{O}), P_i) = 0.31$, $r(E_i(\mathcal{O}), P_i) = 0.43$) [15].

different layers differ significantly. A slightly higher value of correlation is found between centrality at different layers and the centrality of the aggregated network, while the maximum correlation is observed between the values of eigenvector centrality computed on the aggregated topological network and on the aggregated overlapping network.

It is interesting to notice, as shown in Fig. 1.9, that the centrality computed on the aggregated networks (e.g., on the overlapping network) is not correlated with the multiplex participation coefficient of the nodes. In fact, if we fix the value of $E_i(\mathcal{O})$, we observe a large heterogeneity in the values of P_i, and vice-versa.

A number of different approaches have been suggested to define and compute a synthetic measure of node centrality in a multiplex network. A first possibility consists in defining the multiplex centrality as a combination of the centrality scores of each node at the different layers. For instance, starting from a centrality vector of node i, $\boldsymbol{c}_i = \{c_i^{[1]}, \ldots, c_i^{[M]}\}$ (where c represents one of the possible measure of centrality, for instance eigenvector centrality), one can try to condense the information into a single scalar variable, as we did for the degree, in which the role of a node explicitly depends on the structure of the multiplex at all layers. For instance, the authors of Ref. [52] suggested to compute the eigenvector centrality of nodes on each layer α as the normalised eigenvector relative to the largest eigenvalue of

$$\tilde{A}^{[\alpha]} = \sum_{\beta=1}^{M} i^{[\alpha,\beta]} A^{[\beta]} \tag{1.14}$$

where $I = \{i^{[\alpha,\beta]}\}$ is a given influence matrix which determines how the centrality of layer α depends on the structure of layer β.

We now focus on a similar approach, presented in Ref. [15], where the authors studied the contributions of the different layers to the centrality of the nodes for the case of the Terrorists network. Given a two-layer multiplex network (a duplex) and the corresponding adjacency matrices $A^{[1]}$ and $A^{[2]}$, we can construct the following adjacency matrix:

$$\mathcal{M}(b) = bA^{[1]} + (1-b)A^{[2]}, \tag{1.15}$$

which is a convex combination of $A^{[1]}$ and $A^{[2]}$ where b is a parameter taking values in the interval $[0, 1]$. We call such matrix the multi-adjacency matrix. Notice that the parameter b sets the relative contribution of each layer to the multiplex structure. In fact, if $b = 0$ (respectively, $b = 1$) the multi-adjacency matrix of the duplex reduces to $A^{[2]}$ (respectively $A^{[1]}$). We can consider $b = 0.5$ as the benchmark case, where the two layers are given the same weight. Notably, we have $\mathcal{M}(b = 0.5) = \mathcal{O}/2$, i.e. for $b = 0.5$ the multi-adjacency matrix is proportional to the aggregated overlapping network.

For each value of b, $\mathcal{M}$ is a square matrix with non-negative entries. Thus, being satisfied all the hypotheses of the Perron-Frobenius theorem, we can calculate the eigenvector centrality of $\mathcal{M}$ as a function of b. In order to assess the role of each layer in determining the multiplex centrality, we follow this approach: we compute the eigenvector centrality of the benchmark case $b = 0.5$ (corresponding to matrix $\mathcal{O}$); we then compute the eigenvector centrality of $\mathcal{M}$ for a generic value of $\overline{b}$, and we evaluate the Kendall correlation coefficient τ_k between the centrality ranking obtained for $b = \overline{b}$ and the benchmark case $b = 0.5$. Since the multiplex network of the Indonesian terrorists has three layers, we can construct three different duplex networks. The results are shown in Fig. 1.10, where we plot the Kendall coefficient τ_k as a function of b.

As expected, the three duplex have a peak $\tau_k = 1$ for $b = 0.5$. By comparing the three curves we can deduce that T and O have a similar role in determining the centrality of the multi-layer system, in both cases stronger than layer C. In fact, the slopes of the curves, as well as their symmetry/asymmetry, and the symmetry/asymmetry of the extreme cases $b = 0$ and $b = 1$, tell us about the interplay between the two layers in determining the centrality of the multi-layer system. The curve corresponding to the duplex T-O is quite symmetrical, indicating that the effect of T and O on the centrality is very similar. Conversely, the curves corresponding to T-C and O-C are asymmetrical. This means that both layers T and O dominate layer C in determining the centrality of the nodes. If we focus on the case $b = 0$, we obtain three similar values of τ_k. Instead, the three curves display different behavior in the range $0 \leq b \leq 0.5$. In particular, the solid blue

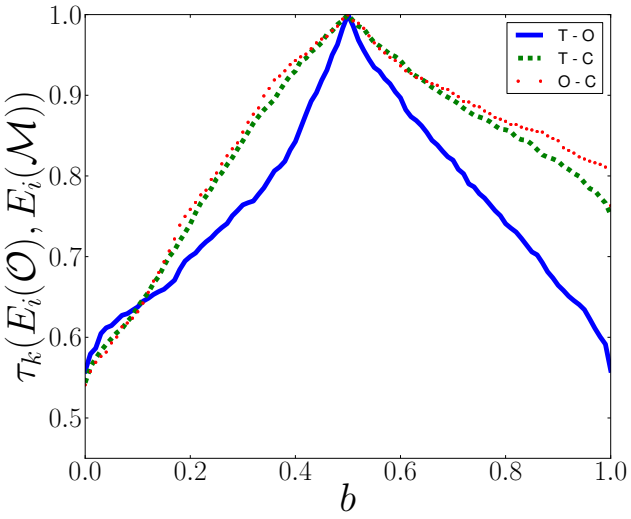


Figure 1.10: For each couple of layers (duplex) of the multiplex network of Indonesian terrorists we plot the Kendall correlation coefficient τ_k between the eigenvector centrality of the benchmark case ($b = 0.5$, i.e. equal weights on both layers) and the generic case of Eq. 1.15 [15].

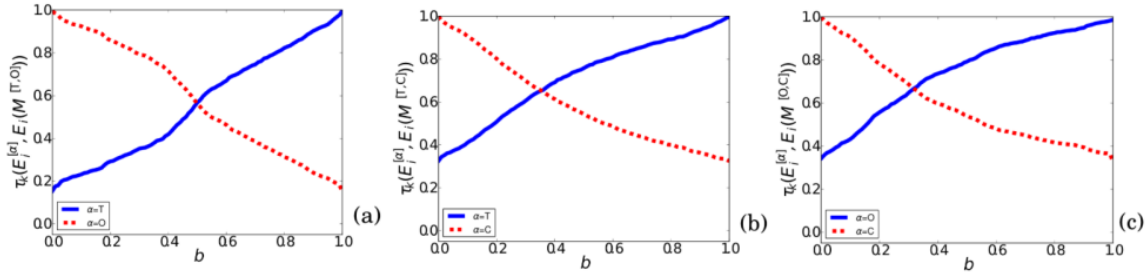


Figure 1.11: As subsets of the original overlapping network, we consider the three duplex $M^{[\mathrm{T,O}]} = bA^{\mathrm{T}} + (1-b)A^{[\mathrm{O}]}$, $M^{[\mathrm{T,C}]} = bA^{\mathrm{T}} + (1-b)A^{[\mathrm{C}]}$ and $M^{[\mathrm{O,C}]} = bA^{\mathrm{O}} + (1-b)A^{[\mathrm{C}]}$. For each possible duplex, we report the Kendall coefficient τ_k between the centrality of each single layer and the corresponding $\mathcal{M}$ as a function of b [15].

curve shows the steepest decrease from the peak (this is also true for $b \geq 0.5$), indicating that layers T and O are more different than layers T and C or layers O and C. For this reason, a small perturbation of the coefficients of $\mathcal{M}$ from the benchmark case affects the centrality of the multi-layer system more for the duplex T-O than for the duplexes T-C and O-C. The largest dissimilarity of the pair T-O is also confirmed by the smallest value of τ_k found for the couple $E_i^{[\mathrm{T}]}$ and $E_i^{[\mathrm{O}]}$, as shown in Fig. 1.8(d).

A slightly different approach provides useful insights about the distribution of cen-

trality in the system under study. Given the three duplex networks, for each one of them we can compute the Kendall coefficient τ_k between the values of centrality obtained for $\mathcal{M}$ and different values of b, and those obtained for each single layer. Results are shown in Fig. 1.11. We note that the value of $\tau_k(E_i^{[\alpha]}, E_i(\mathcal{M}(b = 0.5)))$ in each panel of Fig. 1.11 is equal, respectively, to the value of $\tau_k(E_i(\mathcal{O}), E_i(\mathcal{M}(b = 1)))$ for $\alpha = 1$ and to $\tau_k(E_i(\mathcal{O}), E_i(\mathcal{M}(b = 0)))$ for $\alpha = 2$ on the corresponding curve in Fig. 1.10. In Fig. 1.11(a) the two curves are quite symmetrical and intersect around $b = 0.5$, indicating that the contributions of layers T and O to centrality is similar. Conversely, for both T-C and O-C (respectively, Fig. 1.11(b) and Fig. 1.11(c)) the two curves are asymmetrical and intersect at $0.35 < b < 0.40$, indicating that both layer T and O have stronger impact on centrality than layer C.

These results indicate that multi-layer systems are characterised by non-trivial organisation also with respect to centrality. We conclude this Section by noticing that such definition of multiplex centrality can be easily generalised the to a system of M levels by constructing the adjacency matrix:

$$\mathcal{M} = b_1 A^{[1]} + b_2 A^{[2]} + ... + b_M A^{[M]} \tag{1.16}$$

with the condition that $\sum_{i=1}^{M} b_i = 1$. Once again the benchmark case obtained by fixing $b_1 = ... = b_M = \frac{1}{M}$ coincides with the aggregated overlapping network.

We remark that a different entire class of node centrality measures can be defined by using the properties of random walks on multiplex networks [53]. A particularly interesting example is that of multiplex PageRank centrality proposed in Ref. [54]. The authors of Ref. [54] considered the case of a two-layer multiplex network and defined the multiplex PageRank of the nodes in layer $\alpha = 2$ as a function of the PageRank scores of the nodes in layer $\alpha = 1$. The main idea of this genuinely multiplex measure is that, especially in social systems, nodes can leverage their centrality in one context, such as personal relationships (represented by layer 1) to gain centrality in another context, e.g. professional relationships (represented by layer 2).

Finally, if one represents a multiplex network using the order-4 tensor, an entire class of centrality measures can be obtained as natural extensions to adjacency tensors of the corresponding measures defined on adjacency matrices [56]. For instance, the eigenvector centrality of a node in this formalism can be computed by considering either the eigenvectors of the order-4 tensorial representation of the multiplex or the eigenvectors of the associated supra-adjacency matrix. An interesting application of this class of measures, described in Ref. [55, 56], allows to define the *versatility* of nodes, assigning higher centrality scores to those nodes which act as bridges among different layers.

1.6 Discussion

In this chapter we focused on the main actors of networks, their nodes, and discovered how the allocation of links and resources across the different layers of a multiplex networks impact basic properties such as their activity, degree, clustering, reachability and centrality. Empirical analysis of real-world multiplex datasets, such as transportation networks, co-authorship networks and social networks suggested that nodes take advantage of the layered structure diversifying their relationship, and possibly to perform different tasks. We eventually proposed and reviewed a number of genuinely multiplex descriptors that, compressing information from the vectors associated to a given property into synthetic and meaningful scalar variables, are able to provide an effective description of this additional level of richness pervading the node of the system. In particular, rather than simply generalising standard network measures to the case of multilayer systems, we attempted to create tools able to quantify the added value of multiplexity at the local and global scales. In the next chapter, we change our perspective from nodes to layers and edges, investigating similarities and overlap of the connections across the different levels of a multiplex network.

Chapter 2

Layer and edge properties

In the previous chapter we unveiled how to measure multiplexity from the perspective of the units of our system. Let us consider again the graphical representation of a multiplex network, as shown in Figure 0.2. Investigating node properties, for instance how a node is connected, and how much it is clustered, reachable and central in a network, can be considered a vertical analysis of a multilayer system. Similarly, for multiplex networks it can be of interest to slice the system horizontally and study it from the perspective of the layers. In this Chapter we show how to characterise the structure and similarity between the layers of a multiplex network, each one of those describing the pattern of connections of links of a specific type. Indeed, multiplexity is fundamentally a link property, with the same pair of nodes that can possibly be connected at different levels of the system, giving rise to different levels of overlap. These multi-links can also be considered the most simple multi-layer motifs, which will be extensively discussed in the next Chapter.

2.1 Layer activity and pairwise similarity

Similarly to the case of node activity, it is possible to define the *activity*-vector of each layer α [22] as

$$\boldsymbol{d}^{[\alpha]} = \{b_1^{[\alpha]}, \ldots, b_N^{[\alpha]}\}, \tag{2.1}$$

where $b_i^{[\alpha]} = 1$ if $k_i^{[\alpha]} > 0$, and $b_i^{[\alpha]} = 0$ otherwise. For each layer α, the total layer activity $N^{[\alpha]} = \sum_{i=1}^{N} b_i^{[\alpha]}$ describes the total number of nodes with at least one connection in layer α, with $0 \leq N^{[\alpha]} \leq N$.

The similarity between the activity-vectors of two layers α and β can be measured by mean of the *pairwise multiplexity* [22], which accounts for the fraction of nodes of the

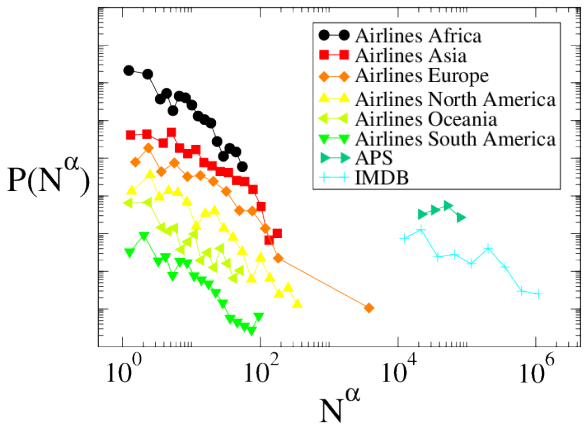


Figure 2.1: Distribution of layer-activity for the continental airline networks, APS and IMDb. A clear power-law trend for $P(N^{\alpha})$ appears in the multiplex transportation systems of continental airlines, which consist of many layers ($O(10^2)$). A weaker, but still heterogeneous behaviour is also observed for IMDb, made of 28 layers only [22].

multiplex which are active on both layers:

$$Q^{[\alpha,\beta]} = \frac{1}{N}\sum_{i=1}^{N} b_i^{[\alpha]}.b_i^{[\beta]}. \tag{2.2}$$

In general $0 \leq Q^{[\alpha,\beta]} \leq 1$, with $Q^{[\alpha,\beta]} = 1$ when all nodes are active in both layers, and $Q^{[\alpha,\beta]} = 0$ when no node is active on both layers. The similarity among the patterns of activity in two layers can also be measured through the *Hamming distance* [22]

$$H^{[\alpha,\beta]} = \frac{\sum_i b_i^{[\alpha]}(1-b_i^{[\beta]}) + b_i^{[\beta]}(1-b_i^{[\alpha]})}{\min(\mathrm{N}^{[\alpha]} + \mathrm{N}^{[\beta]}, \mathrm{N})}. \tag{2.3}$$

where $H^{[\alpha,\beta]} = 0$ if $\boldsymbol{d}^{[\alpha]} = \boldsymbol{d}^{[\beta]}$ and $H^{[\alpha,\beta]} = 1$ if all active nodes are active in no more than one layer. It has been suggested that real multiplex networks are normally characterised by heterogeneous distributions of layer activity and of pairwise multiplexity [22].

Interestingly, it has also been found that many real-world multiplex networks not only are far from being random combinations of the different layers, but their structures rather appear to be determined by hidden geometric correlations [57].

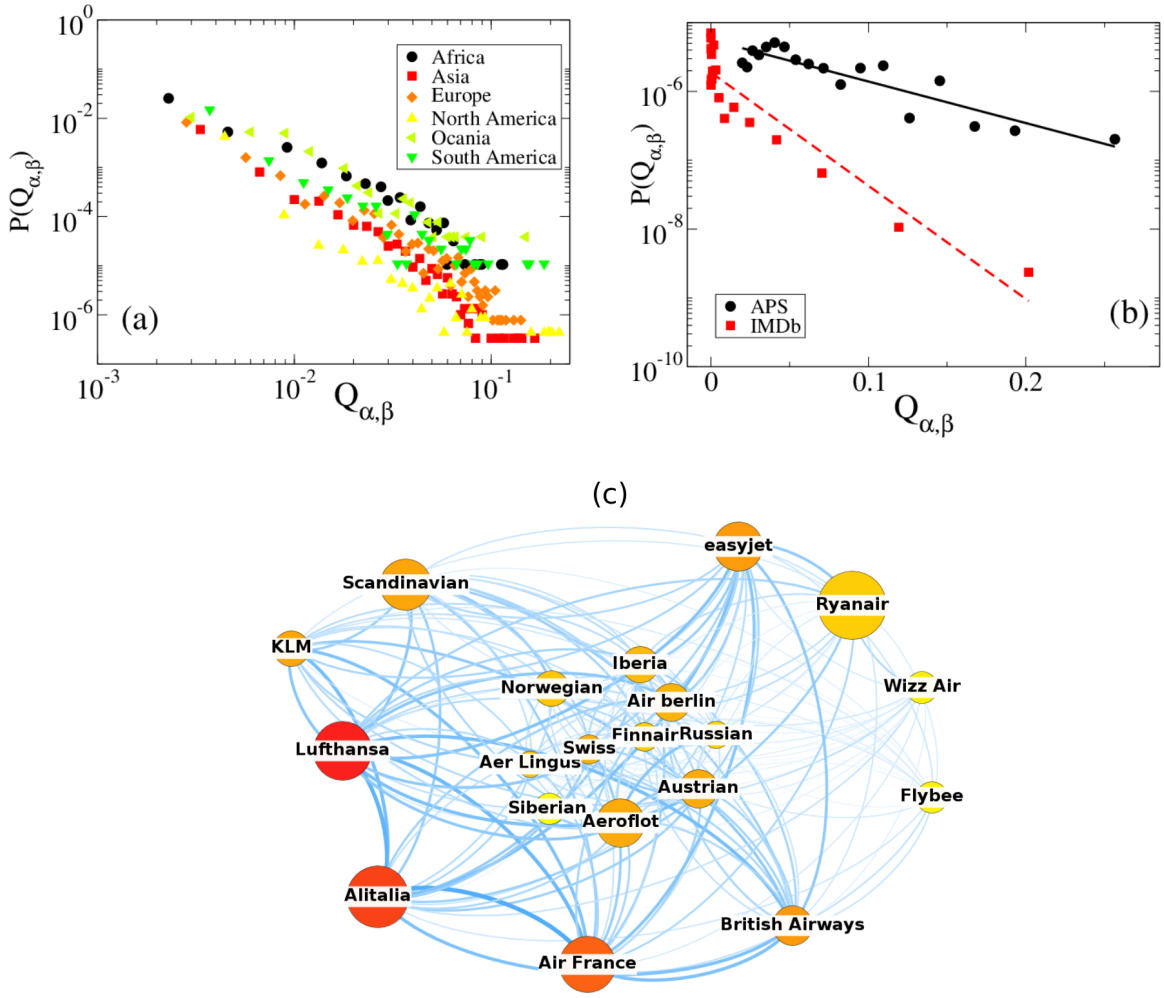


Figure 2.2: The pairwise multiplexity has a power-law behavior in (a) airline networks, while it is exponential in (b) APS and IMDb. In panel (c) we report a graph of the first 20 airlines in Europe by number of covered airports. Each node of this graph represents a layer of the original multiplex network, while the weight of the edge connecting two nodes is proportional to the fraction of nodes present in both layers. The size of a node is proportional to the number of airports in which the corresponding company operates, while the color (from yellow to red) corresponds to the node strength, which in this case is proportional to the total node overlap with other airlines [22].

2.2 Inter-layer degree correlations

Another interesting property observed in real multiplex networks is the presence of correlations between the degrees of the same node at different layers. This is normally signalled by the fact that the probability $P(k^{[\alpha]} = k_1, k^{[\beta]} = k_2)$ to find a node with degree k_1 on layer α and degree k_2 on layer β does not factorise in the product $P^{[\alpha]}(k)P^{[\beta]}(k)$ of the degree distributions of the two layers. In general, given two layers α and β and a generic node property $\boldsymbol{\xi}_i$, the correlation between $\xi_i^{[\alpha]}$ and $\xi_i^{[\beta]}$ can be computed using the *rank correlation coefficient* [22]:

$$\rho^{[\alpha,\beta]} = \frac{\sum_i \left(R_i^{[\alpha]} - \overline{R^{[\alpha]}}\right)\left(R_i^{[\beta]} - \overline{R^{[\beta]}}\right)}{\sqrt{\sum_i \left(R_i^{[\alpha]} - \overline{R^{[\alpha]}}\right)^2 \sum_j \left(R_j^{[\beta]} - \overline{R^{[\beta]}}\right)^2}} \tag{2.4}$$

where $R_i^{[\alpha]}$ is the rank of node i at layer α induced by the property $\boldsymbol{\xi}$, and $\overline{R^{[\alpha]}} = \frac{1}{N}\sum_i R_i^{[\alpha]}$ is the average rank.

The authors of Ref. [58] proposed to quantify inter-layer degree correlations by using the *pairwise mutual information between the degree sequences* of the two layers:

$$I^{[\alpha,\beta]} = \sum_{k^{[\alpha]}} \sum_{k^{[\beta]}} P(k^{[\alpha]}, k^{[\beta]}) \log \frac{P(k^{[\alpha]}, k^{[\beta]})}{P(k^{[\alpha]})P(k^{[\beta]})} \tag{2.5}$$

which is maximal when the degree sequences $\left\{k_i^{[\alpha]}\right\}$ and $\left\{k_i^{[\beta]}\right\}$ are perfectly correlated (or perfectly anti-correlated), and minimal when they are uncorrelated. We notice that a similar set of quantities to measure the inter-layer assortativity has been defined for the order-4 tensorial formulation in Ref. [59]. The same formalism has also been used to analyse the spectral properties of multiplex networks [60, 61].

An alternative approach it is based on defining the quantity:

$$\overline{k^{[\beta]}}(k^{[\alpha]}) = \sum_{k^{[\beta]}} k^{[\beta]} P(k^{[\beta]}|k^{[\alpha]}) \tag{2.6}$$

that is the average degree at layer β of a node having degree $k^{[\alpha]}$ at layer α, and is the multiplex homologous of the nearest-neighbours average degree function $k_{nn}(k)$ traditionally used to quantify degree-degree correlations in single-layer graphs [62]. An increasing (decreasing) trend in $\overline{k^{[\beta]}}(k^{[\alpha]})$ will signal the presence of positive (negative) inter-layer degree correlations between layer α and layer β.

We note that a different type of degree correlations are often studied in single-layer

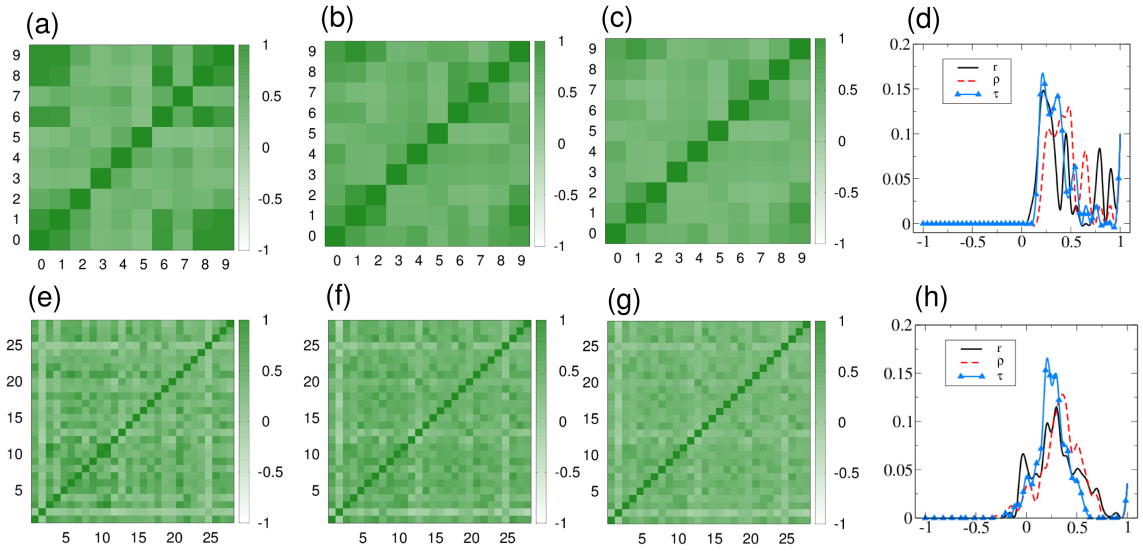


Figure 2.3: Different degree correlation coefficients, namely (a) Pearson's r, (b) Spearman's ρ and (c) Kendall's τ for different couples of layers, and the corresponding distributions (d) are reported for the APS, and show that inter-layer correlations in this system tend to be assortative. A similar pattern is observed in IMDb (panels (e)-(h)). However, some movie genres, like Adult and Talk-Show (respectively corresponding to layer number 2 and number 25 in the diagram) have marked negative inter-layer correlations with almost all the other layers [22].

networks, which we call here intra-layer degree correlations. In particular, the quantity of interest is typically is the average degree of the nearest neighbours (nn) on layer α of a node with given degree $k^{[\alpha]}$ on that layer, i.e. $\langle k_{nn}^{[\alpha]}(k^{[\alpha]})\rangle$. In particular, $\langle k_{nn}^{[\alpha]}\rangle$ is obtained as an average of $k_{nn,i}^{[\alpha]}$ over all nodes with the same degree $k^{[\alpha]}$. The node term can be computed as $k_{nn,i}^{[\alpha]} = \frac{\sum_{j\neq i} a_{ij}^{[\alpha]} k_j^{[\alpha]}}{k_i^{[\alpha]}}$, where $a_{ij}^{[\alpha]}$ are the entries of the adjacency matrix at layer α. If $\langle k_{nn}^{[\alpha]}(k^{[\alpha]})\rangle$ is an increasing (decreasing) function of $k^{[\alpha]}$ if assortative (disassortative) intra-layer degree correlations are present.

Similarly, it is possible to measure the presence of mixed correlations through the function $\langle k_{nn}^{[\beta,\alpha]}(k^{[\alpha]})\rangle$, that is the average degree on layer β of the nearest neighbours on layer β of a node with degree $k^{[\alpha]}$ on layer α. In analogy with the case of intra-layer correlations, the node term is $k_{nn,i}^{[\beta,\alpha]} = \frac{\sum_{j\neq i} a_{ij}^{[\beta]} k_j^{[\beta]}}{k_i^{[\alpha]}}$. We remark here that there exists another possible definition of mixed correlations coefficient, which considers the nearest neighbours of a node on layer α rather then β. Mixed degree investigations have

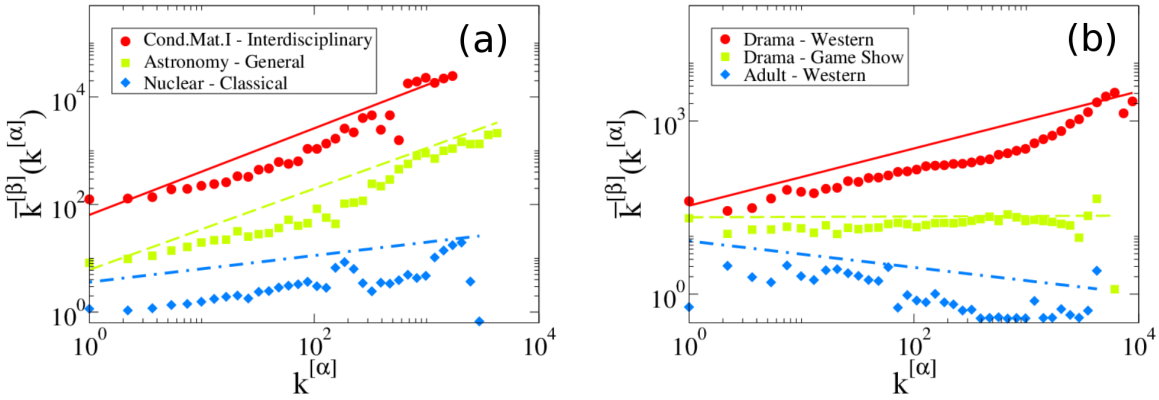


Figure 2.4: The inter-layer pairwise degree correlation function $\overline{k^{[\beta]}}(k^{[\alpha]})$ is shown for various couples of layers α and β, respectively, in (a) APS and (b) IMDb. The lines reported are fit obtained by a power law of the form $\overline{k^{[\beta]}}(k^{[\alpha]}) \sim (k^{[\alpha]})^{\mu}$. in APS only assortative inter-layer degree correlations are observed. All types of inter-layer correlations are instead observed among the layers of IMDb. As an example of positively correlated genres in the IMDb we report the couple Drama-Western. The couple Adult-Western is instead negatively correlated, while Drama Movies are not correlated with Game Show [22].

been investigated both on synthetic and real-world systems, such as the APS and IMDb collaboration networks [63].

2.3 Edge overlap

Multiplexity is fundamentally a link property. Due to the presence of multiple layers, indeed, a pair of nodes (i, j) can be connected through several edges. Given two layers α and β, the *edge overlap* of the pair (i, j) [14, 15] is defined as

$$o_{ij}^{[\alpha,\beta]} = \frac{a_{ij}^{[\alpha]} + a_{ij}^{[\beta]}}{2}, \tag{2.7}$$

where $o_{ij}^{[\alpha,\beta]} = 1$ if i and j are connected at both layers, $o_{ij}^{[\alpha,\beta]} = 1/2$ if they are connected at one layer only, and $o_{ij}^{[\alpha,\beta]} = 0$ if the two nodes are node connected. For a generic number of layers M, the edge overlap is defined as

$$o_{ij} = \frac{1}{M} \sum_{\alpha} a_{ij}^{[\alpha]}. \tag{2.8}$$

This measure can be easily extended to the whole network as

$$o = \frac{2}{N(N-1)} \sum_{i,j \neq i} o_{ij}, \tag{2.9}$$

where the average is computed over all possible pairs of nodes [15], or instead as

$$\omega = \frac{\sum_i \sum_{j>i} a_{ij}^{[\alpha]}}{M \sum_i \sum_j 1 - \delta_{0,\sum_\alpha a_{ij}^{[\alpha]}}} \tag{2.10}$$

where the average is restricted to the pairs of nodes which share at least one edge [58]. Alternative definitions for the local edge overlap on a node i and the total overlap of two layers are suggested in [14] and respectively read

$$\tilde{o}_i^{[\alpha,\beta]} = \sum_{j=1}^{N} \tilde{o}_{ij}^{[\alpha,\beta]} = \sum_{j=1}^{N} a_{ij}^{[\alpha]} a_{ij}^{[\beta]}, \tag{2.11}$$

and

$$\tilde{o}^{[\alpha,\beta]} = \sum_{i<j} a_{ij}^{[\alpha]} a_{ij}^{[\beta]}, \tag{2.12}$$

where $\tilde{o}_{ij}^{[\alpha,\beta]} = 1$ when both layers have a link between i and j and $\tilde{o}_{ij}^{[\alpha,\beta]} = 0$ otherwise. We note that in the same spirit, a similar measure of edge correlations is the so-called *multiplexity* [64], defined as

$$m^{[\alpha,\beta]} = \frac{2 \sum_{i<j} \min(a_{ij}^{[\alpha]} a_{ij}^{[\beta]})}{K^{[\alpha]} + K^{[\beta]}} \tag{2.13}$$

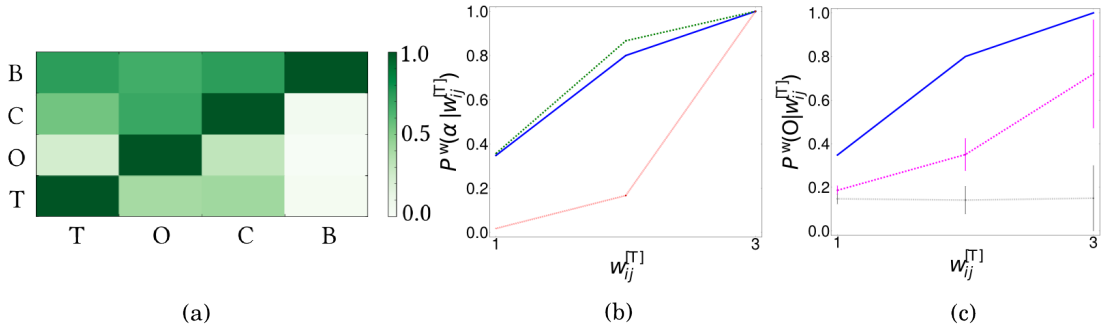
where $K^{[\alpha]}$ ($K^{[\beta]}$) is the total number of edges at layer α (β). Notice that $m^{[\alpha,\beta]}$ takes values in the range $[0, 1]$.

On the same line, a somehow dual quantity is the *edge intersection index*,

$$INT = M \frac{\sum_{i,j=1}^{N} \min_\alpha \left\{ a_{ij}^{[1]}, a_{ij}^{[2]}, \ldots, a_{ij}^{[M]} \right\}}{\sum_{\alpha=1}^{M} \sum_{i,j=1}^{N} a_{ij}^{[\alpha]}} \tag{2.14}$$

which measures the probability of finding a pair of nodes that is connected by an edge on all the M layers of the multiplex [23].

As each layer of a multiplex network is typically sparse, random models for multilayer systems produce very low values of overlap. Conversely, real-world systems are usually characterised by high edge overlap. Let us consider for instance the multiplex formed by all the four layers of the Noordin Indonesian Terrorist Network, i.e. the trust, operational,



(a) (b) (c)

Figure 2.5: (a) For each layer α, we show in the color map the fraction of edges which is also present in each other layer α'. (b) Probability $P^{\rm w}$ of finding a certain link at layer O (solid blue line), C (dotted green line)and B (dashed red line), conditional to the weight $w_{ij}^{[\rm T]}$ of the same link at layer T. (c) The values of $P^{\rm w}$ computed on real data for layer O (solid blue line) are compared to those obtained by randomizing the edges keeping fixed the total number of links $K^{[\rm O]}$ (dashed grey line) or the degree distribution $P(k^{[\rm O]})$ (dotted magenta line) for the operational layer [15].

communication and business layers, so that $1 \leq o_{ij} \leq 4$ for all possible pairs of nodes connected by at least one edge. If we look at the distribution of o_{ij}, we see that 46% of the edges exist in just one of the four layers, 27% are present in two layers, 23% exist in three layers and only 4% are present in all the four layers.

2.4 Reinforcement

Besides the distribution of o_{ij} gives some information about the existence of edge correlations, it is not able to disentangle the relevance of single layers. A slightly more sophisticated quantity we can look at is the conditional probability of finding a link at layer α' given the presence of an edge between the same nodes at layer α:

$$P(a_{ij}^{[\alpha']}|a_{ij}^{[\alpha]}) = \frac{\sum_{ij} a_{ij}^{[\alpha']} a_{ij}^{[\alpha]}}{\sum_{ij} a_{ij}^{[\alpha]}} \tag{2.15}$$

The denominator of Eq. (2.15) is equal to the number $K^{[\alpha]}$ of edges at layer α, while the numerator is equal to the number of such edges which are also present at the layer α'. The conditional probability $P(a_{ij}^{[\alpha']}|a_{ij}^{[\alpha]})$ is shown as a heat-map in Fig. 2.5(a) for the four layers. For instance, the first column shows with a color-code the probability to find a link on layer T given its existence on layer B, C, O or T (obviously, we have

$P(a_{ij}^{[T]}|a_{ij}^{[T]}) = 1$), while the last row represents the fraction of edges in layer T which also exist in layer T, O, C and B. Since layers T and O have a composite internal structure of four levels each, which allows us to assign a weight $w_{ij}^{[\mathrm{T}]}$ and $w_{ij}^{[\mathrm{O}]}$ to each pair of connected nodes i and j, it is interesting to study the probability $P^{\mathrm{w}}(a_{ij}^{[\alpha']}|w_{ij}^{[\alpha]})$ of having a link on layer α' given its weight on the leading layer α, with α corresponding to layers O, and T. In Fig. 2.5(b) we plot the probability of finding a link at layer O, C and B, given the weight $w_{ij}^{[\mathrm{T}]}$ of the link at layer T. Even though in principle $w_{ij}^{[\mathrm{T}]} = 4$ is possible, none of the edges appears together in all classmates, friendship, kinship and soul-mates sub-layers of the trust layer. In all the three cases, P^{w} is an increasing function of $w_{ij}^{[\mathrm{T}]}$. Fig. 2.5(b) suggests that the stronger the trust connection between two terrorists the higher the probability for them to operate together, communicate or having common business. In particular, for layer O and C, which are the ones that have a number of nodes comparable to the one of layer T, already a value of $w_{ij}^{[\mathrm{T}]} = 2$ implies that the two people have common operations and communications in 80% of the cases. If $w_{ij}^{[\mathrm{T}]} = 3$, then the probability that the edge $i - j$ exists in all the three remaining layer is equal to 1.

This phenomenon can be explained in terms of social *reinforcement* [65], meaning that the existence of strong connections in the Trust layer, which represents the strongest relationships between two people, actually fosters the creation of links in other layers and produces a measurable effect on the probability to operate, communicate and do business together. Despite we do not have longitudinal information to test the hypothesis that original trust connections actually caused the creation of links in other layers by means of social reinforcement, in this particular case we have to stress that the strength of the trust relationship between two individuals is higher if they had been kin, classmates, soul-mates, and/or friends, respectively. This means that, with high probability, the establishment of any of the four Trust relationships between i and j preceded by several years the establishment of any communication, operational or business relationship registered during the collection of the data set. Consequently, it is not too pretentious to suggest that a social reinforcement mechanism took place in this small social system, and that trust relationships have actually *caused* the subsequent communication and the collaboration among the terrorists.

In order to statistically validate these results, in Fig. 2.5(c) we report the expected values of P^{w} obtained by randomizing the non-leading layers while keeping fixed either the total number of links $k^{[\alpha]}$ or the degree distribution $P(k^{[\alpha]})$. In the first case, each non-leading layer is an Erdös-Renyi random graph and P^{w} is not even correlated with the weights on layer T, as expected. In the second case, which is an extension to multiplexes of the configuration model, for each weight $w_{ij}^{[\mathrm{T}]}$ the conditional probability to

find an edge on the operational layer is systematically lower in the randomized networks than in the original one. Hence, we can conclude that inter-layer correlations among the heterogeneous degree distributions of the various levels do not provide an ultimate explanations to the founded results for P^{w} and that the Trust layer is genuinely *driving* the observed connection pattern. Results analogous to that of layer O, were also found for layers C and B.

Similar results are obtained considering the operational network (instead of the trust network) as leading layer, but in this case the conditional probability of finding an edge in T, C and B given its weight in O was substantially smaller than those reported in Fig. 2.5(b) and Fig. 2.5)(c)(figure not shown). This is not surprising at all, since while it is clear that a stronger level of trust between two individuals can boost their communications and their common operations, we expect a weaker causality between the strength of different operations two individuals have shared and their trust and communications. The existence of a weaker interaction between the operational layer and the other three layers increases the validity of our hypothesis that the trust layer is indeed controlling the overall structure of the multiplex network through a social reinforcement mechanism and that the relative importance of the trust layer for the formation of edges on other layers is not a mere consequence of existence of sub-layers.

2.5 Discussion

In this chapter we saw how to study the structure and similarity between the layers of interactions in a multiplex network, and how to unveil and measure multiplexity at the edge level. Real-world networks are characterised by high levels of overlap and the existence of non-trivial correlations among the different levels at which interactions can occur. Still, a fundamental issue remains somehow open, and concern assessing whether the presence of more than one interaction layer indeed provides more information about the structure of a system compared to a classical single-layer network representation. In many cases, aggregated networks such as $\mathcal{A}$ and $\mathcal{O}$ appear to provide an inadequate description of the system. At the same time, keeping track of all the information concerning the different nature of the links leads to longer and more complicated algorithm for the analysis of our systems and requires larger memory. Hence, is it possible to quantify how much information is lost when we aggregate some or all the layers of a multiplex network to obtain a lower-dimensional representation? What is in general the best representation of a multiplex network which maximise the trade-off between minimising information loss in the aggregation procedure still preserving the layered nature of the system? The authors of Ref. [23] tackled the problem of multiplex reducibility by

drawing on an existing formal parallel between density operators of quantum systems and Laplacian matrices of graphs, and extending the concept of Von Neumann entropy of a graph to the case of multiplex networks. They proposed a greedy procedure, based on the estimation of the quantum Jensen-Shannon divergence between layers, which allows to successively aggregate the most redundant layers of a multiplex and to obtain a more compact representation which uses the minimal number of layers while maximising the distinguishability between the multiplex and the single-layer representation of the same system. An interesting result of the paper is that different multilayer systems allow different levels of reducibility, with man-made systems being the least reducible and biological and social systems showing the highest levels of redundancy [23].

Chapter 3

Motif analysis

In the previous chapter we focused on how to measure multiplexity at the layer and edge level in a multiplex system. More generally, many networks are characterised by the presence of non-trivial structures at the microscopic scale. In particular, biological networks are rich in certain subgraphs, because these are crucial for the stability of the system and for the efficient processing of information. Motifs have been largely studied in neuroscience both in networks of anatomical connectivity and in networks of correlations in the functional activity of different brain regions. To shed new lights on the intimate relations between structure and function in the human brain, we consider the two networks as the layers of a multiplex brain network, and we investigate the presence of statistically overabundant subgraphs spanning across the two layers in several healthy subjects. We provide a mathematical framework for multi-layer motif analysis and identify over-represented subgraphs associated to the existence of overlap and structural balance in the two layers, as well as the existence of significant reinforcement mechanisms among the structural and functional connections in the human brain.

3.1 Motifs in brain networks

In neurosciences, it is widely acknowledged that the emergence of several pathological states is accompanied by alterations in brain connectivity patterns [66, 67]. Indeed, empirical studies have lead to the hypothesis that the brain relies on the coordination of a scattered mosaic of distant brain regions, forming non-random a weblike structure of neural assemblies, and that brain dysfunctions are related to a lack of such coordination [68]. In the last decade, the combined use of advanced neuroimaging techniques and of mathematical tools to characterise the structure of a complex network has significantly

improved our understanding of how the brain works. However, it is important to notice that we have two fundamentally different ways to study brain connectivity, since data can reflect either anatomical properties of the brain, or functional neural activity.

In the first case (*anatomical brain networks*) we construct networks whose nodes are usually putative brain regions and the links represent physical connections among them, while in the second case (*functional brain networks*) each node represents an area of the brain, usually consisting of neural assemblies, and each link indicates the presence of a functional interaction between the activity (electrical, magnetic or hemodynamic/metabolic) of two areas.

Despite the evident relations between the two type of networks, comparison between anatomical and functional brain networks is not straightforward [69, 70]. Theoretical studies support the idea that anatomical connections can determine some aspects of brain dynamics [69], but it is less clear how in general the anatomical connectivity supports or facilitates the emergence of the properties of functional networks.

Correspondence between functional and structural networks remains thus an active research area [71–73]. A better understanding of how anatomical connections support communication, correlations and synchronisation of brain activities is necessary to read normal neural processes, as well as to improve the identification and prediction of alterations in brain diseases.

In this chapter we contribute to unveiling the delicate relations between structure and function in the human brain by focusing on *network motif analysis*, a tool that has revealed quite successful in network science. A network motif is a small subgraph that is statistically over-represented in a complex network with respect to a given null model [74]. Empirical evidence suggests that motifs are a key concept from RNA structures to social networks [74–77]. The emergence of motifs, i.e. the abundance of certain types of subgraphs in a given network, seems to be related to the robustness of the system, or to the stability of the dynamical or signalling circuits that each motif represents [74–77].

In previous studies, functional interactions have been found to variate with the patterns of local structural motifs in the monkey cortex [78]. Similarly, functional integration of cortical areas in monkeys seem to be strongly determined by some properties (e.g. density and symmetry) of structural motifs [79]. Neurocomputational modeling indicates, for instance, that neuronal networks motifs might play a role on information transmission delays and on long- and short-term memory [80]. Recent results suggest that network motifs analysis can provide significant new markers for the progression of Alzheimer's disease [81]. Motif analysis has been applied separately, both to anatomical

and functional brain and, although some motifs which are considered central to information processing in the brain have emerged [82–84], the interplay between structural and functional motifs is still not well understood.

Here, we investigate the relation between structure and function in the brain by generalising motif analysis to the case of multiplex networks, and by detecting *multiplex motifs* in the brain. We note that the multi-layer formalism has been recently applied to other brain data such as connectivity matrices estimated from magnetoencephalographic data for getting a more complete picture of neural interaction across different frequency bands [85, 86].

The multiplex motifs we are interested in are small multi-layer connected subgraphs which are statistically overabundant in multiplex networks describing real systems. We note that the related problem of finding isomorphisms in multi-layer networks has been considered in Ref. [87]. The layered organization of triadic connections considered in Chapter 1 [15], as well as other possible generalisation of clustering [49, 88–90] can also be considered as specific families of motifs. In particular, in this Chapter we use multi-layer motif analysis to study multiplex networks with two layers constructed from structural and functional information on the brains of healthy subjects, respectively obtained by Diffusion Tensor Imaging (DTI) and resting-state functional MRI (rs-fMRI) [91]. In these networks, nodes are defined as Regions of Interest of the brain (ROIs). The edges of the structural DTI layer represent the estimated white matter connection strength between any pair of ROIs, while links in the functional network indicate functional correlations between the fMRI time-series of the two corresponding ROIs. As we will show, our approach to detect multiplex anatomical/functional motifs is able to provide useful insights on multiplex networks derived from multiple brain modalities of several subjects [91].

3.2 Mathematical framework

Single-layer motifs. In single-layer networks, standard motif analysis searches for small subgraphs that are statistically over-represented in a given graph G with respect to a null model [74]. In practice, the frequency of each subgraph g in graph G is compared with the expected frequency of that subgraph in an appropriately randomised version of the graph G, e.g. in the family of random graphs having the same number of nodes and the same number of edges of G. If the actual frequency of the subgraph g in in G is larger than that expected in the null model and the difference is statistically significant, then g is an overrepresented subgraph, i.e. a network motif. Small connected subgraphs are

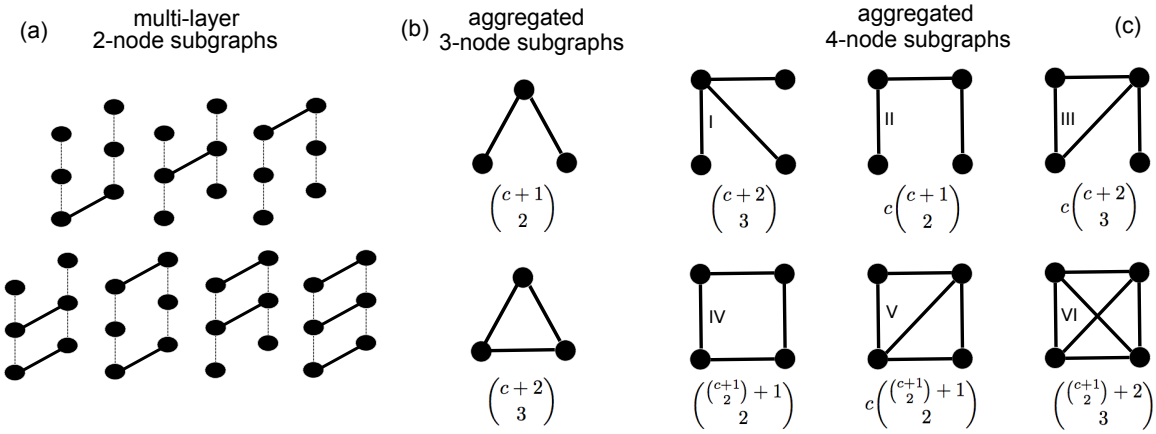


Figure 3.1: A simple edge on the aggregated network can originate from different multi-layer subgraphs. For $M = 3$ layers there exist $c = 7$ different types of multi-links ($s = 1$, see Eq. 3.2) (a). At the aggregated level, we distinguish two subgraphs with $n = 3$ nodes (b), the chain (top) and the clique (bottom), and six motifs with $n = 4$ nodes (c), usually respectively known as the star, the chain and the 3-loop-out (top, left to right), the box, the semi-clique and the clique (bottom, left to right). We report the number of different multi-layer motifs giving rise to each aggregated subgraph as a function of the number of multi-links c [91].

typically classified at two different levels: they are first sorted according to their number of nodes n and then classified based on the number ℓ and placement of their links. If G is undirected, the smallest subgraphs of interest are those with $n = 3$ nodes. In this case, two different connected subgraphs can be identified, namely the chain, also known as triad, and the complete graph, or triangle. We have then six different subgraphs with $n = 4$ nodes, nineteen subgraphs with $n = 5$ nodes, with this number growing fast as the number of nodes increases. Because of the decreasing statistical significance of larger subgraphs and the growing computational cost associated to their detection, motif analysis in real-world networks is usually limited to small subgraphs consisting of a few nodes.

Motifs in multi-layer networks. A multiplex network $\mathcal{M}$ is a convenient representation for systems where nodes are related through different types of interactions. We already saw that any multiplex network can be aggregated into a unique and non-ambiguous aggregated network $\mathcal{A}$. The simplest approach would be then a single-layer motif analysis on network $\mathcal{A}$. While such analysis provides useful information on micro-scale connectivity patterns at the aggregated level, information is lost on the significance of the layered organisation of interactions. A given subgraph g observed in the aggregated graph $\mathcal{A}$ can indeed originate from different combinations of the edges across the

layers of the system, so that we are interested to enumerate the different types of multi-layer motifs contributing to each subgraph in the aggregate network. We suggest here a classification of motifs in multiplex networks on three levels [91]. At the *first level*, multi-layer connected subgraphs are categorised according to their number of nodes n. At the *second level*, they are classified according to the subgraphs they generate in the corresponding aggregated network $\mathcal{A}$. At the *third level*, for each subgraph in $\mathcal{A}$, the different multiplex subgraphs are distinguished by looking at the exact pattern of connections across the different layers.

Multi-layer motifs with $n = 2$ nodes. Due to the richness provided by the presence of more than one layer, the smallest motifs of interest in multiplex networks are already those with $n = 2$ nodes. Such multilayer motifs correspond in practice to the different types of *multi-links* connecting two nodes [14]. An edge between two nodes appears in the aggregated network if the two nodes are connected on at least one of the layers. We now derive the total number of different multi-layer configurations giving rise to a single link in $\mathcal{A}$ [91]. In particular, in a multiplex with M layers we have $\binom{M}{1}$ configurations where an edge between two nodes exists only in one of the layers, $\binom{M}{2}$ configurations such that two nodes are linked at two different layers, and so on. Hence, the total number of distinct multi-links is equal to

$$c = \sum_{m=1}^{M} \binom{M}{m} = 2^M - 1, \tag{3.1}$$

where we have neglected the degenerate configuration in which there is no edge between the two nodes at any layer, since in that case the corresponding subgraph is disconnected. Let us consider for instance a multiplex network with $M = 3$ layers. As shown in Figure 3.1(a), we have in total $c = 7$ possible motifs with $n = 2$ nodes, of which $\binom{3}{1} = 3$ have one edge at a single layer, $\binom{3}{2} = 3$ with edges on two layers and one complete multi-link with edges at all levels $\left(\binom{3}{3} = 1\right)$.

An interesting case is that where the edges at a given layer can be of different types, e.g. signed or coloured edges, and such information is lost in the aggregate graph. Let us denote by s the number of different types of edges allowed on each layer (for instance $s = 2$ for a signed network with positive and negative edges), with s equal on every layer. Since each of the m connections can be of s different kinds, to correctly count all the possibilities the terms $\binom{M}{m}$ must be multiplied by a factor s^m, and the total number of multi-links becomes equal to

$$c = \sum_{m=1}^{M} s^m \binom{M}{m} = (s+1)^M - 1. \tag{3.2}$$

If each layer α has a different number $s^{[\alpha]}$ of edge types, the total number of multi-links is equal to

$$c = \left[\prod_{\alpha=1}^{M} (s^{[\alpha]} + 1) \right] - 1. \tag{3.3}$$

The latter formula correctly reduces to Eq. (3.2) and Eq. (3.1) respectively if $s^{[\alpha]} = s$ $\forall \alpha$, and if $s = 1$.

Multi-layer motifs with $n = 3$ nodes. For motifs with more than $n = 2$ nodes, the problem of counting the number of multi-layer configurations of a given aggregated motif is equivalent to that of finding the number of non-isomorphic ways in which that subgraph can be coloured by choosing edges from c different colours [91].

Let us focus first on the case of connected subgraphs with $n = 3$ nodes. In such case we distinguish two different subgraphs on the aggregated network $\mathcal{A}$, the triad $g_{n=3,\ell=2}$ and the triangle $g_{n=3,\ell=3}$, shown respectively at the top and the bottom of Figure 3.1(b). Let us consider a *triad* formed by $\ell = 2$ specific multi-links, each chosen from the c possible ones. In general there are c^2 ways to colour a labeled triad. However, each possible configuration of two coloured multi-edges generates 2 isomorphic configurations. Hence, the number t of different multi-layer triads is equal to the number of different unordered selections with repetition of $\ell = 2$ multi-links of c types, i.e.

$$t = \binom{c+1}{2} \tag{3.4}$$

Similarly, let us consider a *triangle* formed by $\ell = 3$ specific multi-links. In general there are c^3 to colour a labeled triangle. However, each possible configuration of two coloured multi-edges generates $3! = 6$ isomorphic configurations. Consequently, the number T of multi-triangles is equal to the different unordered selections with repetition of $\ell = 3$ multi-links, i.e.

$$T = \binom{c+2}{3} \tag{3.5}$$

Multi-layer motifs with $n > 3$ nodes. At difference with the multi-layer triads and triangles, larger multiplex subgraphs are in general not characterised by symmetry classes as simple as those of the triad or triangle. As a consequence, counting the number of multi-layer patterns associated to the same aggregated subgraph is in general more complicated. The basic idea is to decompose each structure as a combination of smaller motifs, such as multi-links (whose multiplicity is equal to c), pairs and triples consisting of symmetric multi-links (for which the previous formulas introduced in Eqs. (3.4) and (3.5) apply) [91].

For simplicity, let us focus on the motifs with $n = 4$ nodes shown in Figure 3.1(c). The *star* I is the only motif where all the links are indistinguishable. Hence, it can be generated in a number of multi-layer configurations corresponding again to selecting $\ell = 3$ unordered multi-links with repetition. For the *chain* II, the multi-layer multiplicity can be determined as the product between the number of possible central multi-links and that of the pair of symmetric external edges. The *3-loop-out* III corresponds to the product of the number of multi-links with the number of multiplex triangles. For the *box* IV, we can decompose the problem into selecting in an unordered way with repetition two different pairs, each composed of two symmetric multi-links. The multi-layer multiplicity of the *semi-clique* V is equal to the product of one multi-link and a box. At last, for the multi-layer *clique* VI we can decompose the problem into selecting in an unordered way with repetition three different pairs, each composed of two symmetric multi-links. Similar techniques can be used to compute the multi-layer multiplicity of motifs with $n > 4$ nodes. The exact number of multi-layer configurations corresponding to the different motifs with $n = 3$ or $n = 4$ nodes is reported in Figure 3.1.

3.3 Structural and functional layers in the human brain

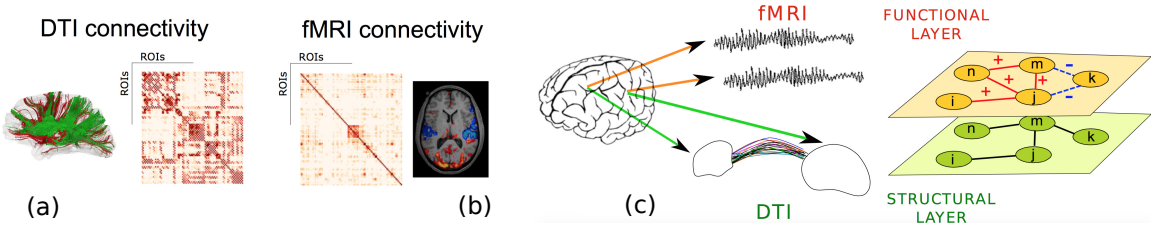


Figure 3.2: DTI connectivity provides the strength of axonal connections among regions of interest (ROIs) of the brain for the anatomical network (a). Positive and negative correlations in the activity of the different areas are measured through fMRI for the functional layer (b). Such information can be combined to construct a multi-layer brain network (c) [91].

We study a data set of multiplex brain networks obtained from the USC Multimodal Connectivity Database (`http://umcd.humanconnectomeproject.org`), an open repository for brain connectivity matrix sharing and analysis [92]. Each multiplex consists of two layers representing the structural (anatomical) and functional connections among the brain areas of the corresponding subject, respectively inferred by means of Diffusion Magnetic Resonance Imaging (DTI) and functional Magnetic Resonance Imaging (fMRI). All the scanning parameters and pre-processing details of both DTI an fMRI connectivity matrices are explained in [93]. The anatomical connectivity networks are based on

the connectivity matrices obtained by DTI data from 19 healthy participants. Whole brain deterministic tractography was performed using the fiber assignment by continuous tracking algorithm. Fiber tractography was carried out by propagating fibers from each voxel with a maximum turn angle of 50° followed by a spline filter smoothing. Each element of the matrix represents thus an approximation of the anatomical strength between the corresponding pair of brain regions. The elements of this matrix gives the estimated number of fibers between different anatomical regions of interest ($N = 264$ nodes in all the networks) spanning the cerebral cortex, subcortical structures, and the cerebellum [94]. The functional connectivity networks are based on the matrices extracted from fMRI recordings as explained in [93]. All fMRI data sets (segments of 6 minutes recorded from the same 19 healthy subjects) were normalized, corrected and sub-sampled from the same set of anatomical regions as for anatomical connectivity. These 264 putative functional regions were shown to more accurately represent the information present in the brain network relative to other voxelwise and atlas-based parcellation approaches. To eliminate low frequency noise (e.g. slow scanner drifts) and higher frequency artifacts from cardiac and respiratory oscillations, time-series were digitally filtered with a finite impulse response filter with zero-phase distortion (bandwidth $= 0.01 - 0.1$ Hz). For the functional connectivity, linear correlation were estimated between time series of each of the 264 brain regions. Correlation coefficient were then variance-stabilized by applying the Fisher's Z-transform in order to generate 264×264 whole brain functional connectivity matrices for each subject.

Consequently, the weight of an edge in the DTI layer indicates the presence (and strength) of axonal connections between the corresponding areas, while the weight of an edge in the fMRI layer is proportional to the correlation in the time-series of hematic flow activity associated to the two areas. The functional connectivity matrices reflect both positive and negative correlations. Anti-phase oscillatory patterns are generally related to different structural factors such as time delay in coupling, modularity and the coupling types (excitatory or inhibitory) [71, 95, 96]. Although the cognitive significance of negative correlations is still speculative, recent evidence suggests biological relevance [97] and potential behavioral significance [98, 99]. The study of negative weights is therefore necessary for understanding brain network organization.

In the following we will refer to $\mathcal{D} = \{d_{ij}\}$ as the physical connectivity matrix, where d_{ij} represents the weight of the physical connection between node i and node j. For each subject such matrix consists of a single connected component and it is sparse ($d_{ij} = 0$ for many pairs $i, j.$). Functional data, conversely, in principle represent a fully connected graph. For such reason we thresholded such graph, keeping only significant functional links (both positive and negative ones). The significance was set at $p < 0.05$, corrected

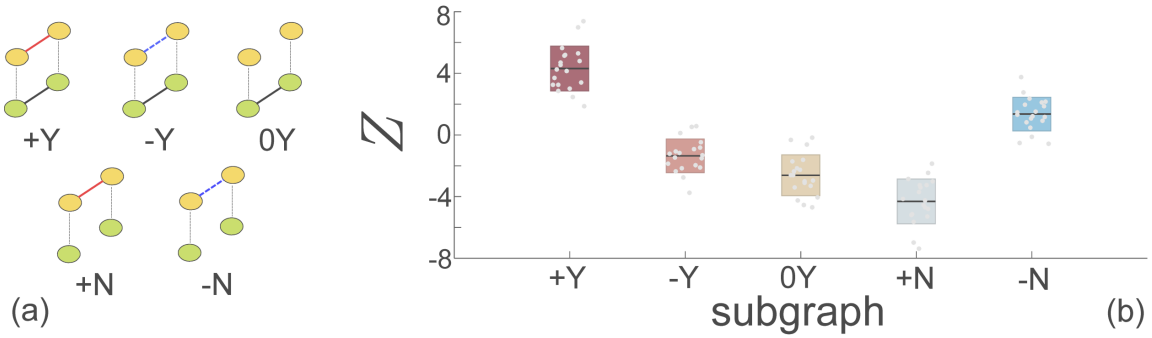


Figure 3.3: The presence (Y) or absence (N) of links in the structural layer $\mathcal{D}$, and the existence of a significant positive link (+), negative link (-) or its absence (0) in the functional layer $\mathcal{F}$ give rise to $c = 5$ (see Eq. 3.3, $s^{[d]} = 1$, $s^{[f]} = 2$) (a). The Z-scores of the subgraphs show that +Y is a strongly over-represented motif, indicating correlation between the co-activation of two ROIs and the existence of a structural link. Conversely, the motif -Y is roughly as likely in real and random data (actually slightly under-represented in real data), suggesting that significant negative functional links appear to be at random relatively to the physical connections (b). Results are shown for 19 healthy subjects [91].

for multiple comparisons [100]. We will refer to such thresholded functional graphs as $\mathcal{F} = \{f_{ij}\}$, where f_{ij} represents the weight of the significant functional connection between node i and node j (if the link is not significant we set $f_{ij} = 0$). We indicate by

$$\mathcal{M} = \{\mathcal{D}, \mathcal{F}\}. \tag{3.6}$$

the multiplex brain network $\mathcal{M}$ encoding information on such structural and functional layers, and illustrated in Figure 3.2. The correlation f_{ij} between the fMRI activity of two ROIs i an j can be either positive (+, red links), or negative (-, blue links) or non-significative (0, no link). Conversely a structural edge between two ROIs might either exist (Y, green links) or not (N, no link).

3.4 Multi-edges in multiplex brain networks

We now move our attention to the investigation of multi-layer motifs in brain networks [91]. We first consider the structural layer $\mathcal{D}$ and the functional layer $\mathcal{F}$ as binary unweighted networks. Since links are signed in the functional layer $\mathcal{F}$, i.e. $s^{[f]} = 2$, and simple in the structural layer $\mathcal{D}$, i.e. $s^{[d]} = 1$, from Eq. (3.3) we have $c = 5$ elementary motifs with $n = 2$, namely +Y , -Y , 0Y , +N , -N. They are illustrated in Figure 3.3(a).

We report in Figure 3.3(b) the Z-score of each multi-link g

$$Z(\mu_g) = \frac{\mu_g - \bar{\mu}_g}{\sigma_g}, \tag{3.7}$$

where the abundance μ_g of the subgraph in the real system is compared with the average abundance $\bar{\mu}_g$ and standard deviation σ_g of what found on a suitable randomised network. A high positive value $Z(\mu_g)$ indicates that g is a significant recurrent motif.

In order to consider a suitable null-model, it is necessary to take into account the division of the brain into two distinct hemispheres. For such a reason, we randomise the links of the layer by performing the following block configuration model. For each node, tot only we preserve its total degree at that layer, but we also keep fixed its number of connections towards regions in the same brain hemisphere and those towards the other hemisphere. Such block configuration model can be practically implemented by performing two standard configuration models for the intra-hemisphere links in the right and in the left hemisphere, and by performing a bipartite configuration model on the inter-hemisphere links. In Figure 3.3(b) results are shown for a multiplex null-model where we kept fixed the signed functional layer and performed a block configuration model on the structural layer, with 100 randomisation for each subject.

Interestingly, the motif +Y (corresponding to the concurrent presence of a positive fMRI correlation and of a direct connection in the DTI layer) is significantly over-represented, while +N (positive correlation and absence of edge) is markedly under-represented. This is a remarkable result as it supports the hypothesis that functional positive links are definitely correlated with the structural network. Conversely the motif -Y is as likely in real data as in the random model, which indicates that two brain areas physically connected do not correlate with negative functional interactions between their dynamics. Results are shown for 19 different subjects.

3.5 Triads and triangles in multiplex brain networks

Having fixed the abundance of the $c = 5$ multi-links, we are interested in knowing the significance of each higher order motif. It is possible to extend such motifs analysis to larger subgraphs, such as motifs with $n = 3$ nodes [91]. According to Eqs. (3.4) and (3.5), in the considered multiplex networks there are $t = 15$ multi-layer triads and $T = 35$ multi-layer triangles. For each subgraph with $n = 3$ we compute now the Z-score by comparing the value found in the real data with the average value and standard deviations of a randomised ensemble of networks, where the unsigned structure

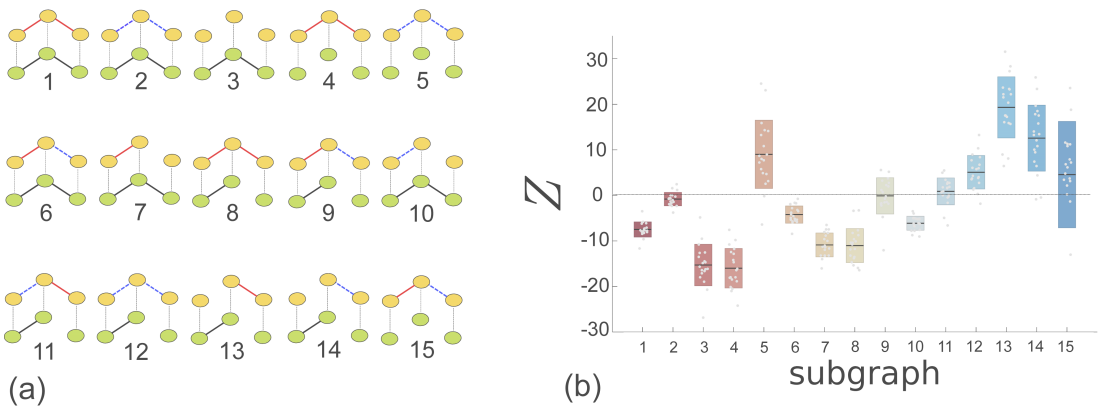


Figure 3.4: The $t = 15$ different multi-layer triads are shown in (a). Their corresponding value of the Z-score are reported in (b). Triad 5, 13 and 14 emerge as statistically validated recurrent motifs ($p < 0.05$, corrected for multiple comparisons) [91].

of the overlapping network is kept fixed, but the different multi-links were shuffled. In Figure 3.4 we see for instance that, even if the multi-link +Y is over-represented, the triad formed by two multi-links of such type is underrepresented, with $Z \approx -8$. This is due to the fact that in the real data, if one region i is connected to two other regions j and k with both a structural and positive functional links, it is very unlikely that j and k are no connected at all. We notice that for the same reason almost all multi-layer triads with at least one multi-link of type +Y have a negative Z-score, as shown in Figure 3.4, with the exception of triad of type 9. In this subgraph, the central node is indeed strongly connected to one of the other region, both physically and with a positive functional link, but only weekly connected to the other with a negative functional link and no structural link. Hence, the two external regions of the triads do not appear to be communicating much, and the resulting Z-score is approximately 0.

We now analyse which other multi-layer triangles are overabundant in the real data and propose possible underlying reasons for this phenomenon. The high Z-score of the first multi-layer triangles in Figure 3.5 confirms that, if one region i is connected to two other regions j and k at both layers, j and k are also directly connected by links at both layers as well. Triangles of type 3 are significative due to high clustering in the structural layer. Triangles of type 4 have a high Z-score because of structural balance (three positive values of correlations) in the functional layer. Similarly, many other multi-layer triangles appear to be overabundant due to the existence of signed balanced triangles in the functional layer (given either by three positive links, as in triangles of type 1, 4 and 18, or by one positive link and two negative link, as for triangles of type 10, 22, 25, 34). The mentioned motifs were confirmed to be statistically over-represented from the reference null model by using a $p < 0.05$ significance level, which control for the

multilayer structure of the graph. A current open problem in the analysis of functional networks is how to effectively eliminate links determined by spurious correlations in the activity (time-series) of different brain regions. We note that issues related to the definition of functional links might be present also here at the multi-layer level, where structural and functional networks are analysed together.

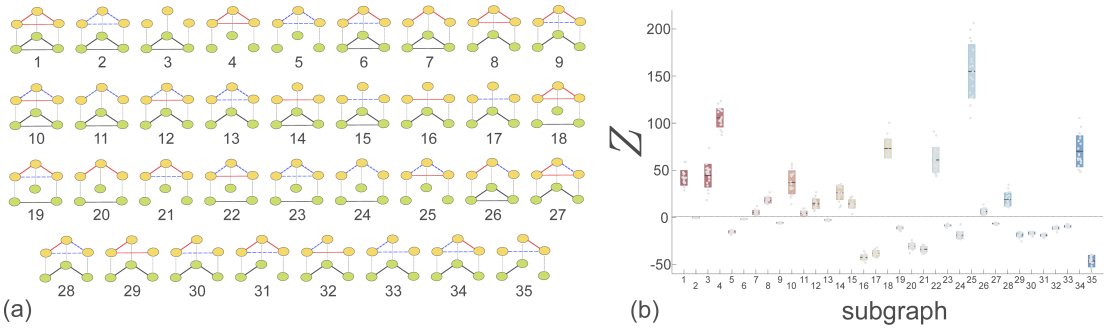


Figure 3.5: The $T = 35$ different multi-layer triangles are shown in (a). Their corresponding value of the Z-score are reported in (b). Triangles 1, 4, 18, 22, 25 and 34 emerge as statistically validated recurrent motifs [91].

3.6 Discussion

The recent prevalence of applications involving multidimensional and multimodal brain data has increased the demand for technical developments in the analysis of such complex data. Modeling the human brain as a complex network has unveiled the presence of charactaristic non-trivial connectivity features (small-worldness, power-law degree distributions and modularity among others) in many structural and functional networks [101]. The interplay between anatomical connections and functional interactions is a current challenge for understanding general brain functioning. In the last decade, some studies have thus directly compared these connectivity structures to better investigate possible direct mappings at the network level [78, 102]. In this study we addressed a fundamental problem in multimodal brain networks analysis: the organisation of the complex mosaic of brain motifs in anatomical and functional connectivity.

By considering multi-layer networks we have identified nonrandom motif structures in multimodal brain networks. In contrast to current approaches, which considers motif analysis on separate brain modalities [82–84], this work provides the first evidence of joint anatomo-functional motifs in human brain networks. In line with previous studies [103, 104], our results confirm the complex relationships between structural connectiv-

ity and coupling of brain dynamics. Significant multi-layer triads differ from those usually obtained from single structural connectivity in both humans [84] or monkeys [78, 82]. This suggests that the multi-layer formalism constitutes could be a very appropriate choice for the analysis of multimodal brain networks.

Our finding of a positive fMRI correlation between brain areas connected by a direct physical link is consistent with prior works [105]. In some cases, this increased functional connectivity can be explained by the spatial proximity of areas [101], but distant regions can also display strongly coherent dynamics without direct physical connections [70]. In agreement with previous works, over-represented multi-layer triads involving negative functional links could also suggest a decrease in the anatomical connectivity that correlates with negative resting state correlations [105]. Nevertheless, more refined neuro-computational models are needed to fully explain the mechanism of this phenomenon.

A limitation of our study is the symmetrical configuration of the analyzed brain connectivity matrices, which is a consequence of the inherent symmetrical properties of DTI techniques (fiber estimation cannot distinguish between afferent and efferent projections) and of the undirected functional interaction obtained by linear correlations (for a network of $N = 264$ nodes, directed interactions could be estimated but from much longer time series [106]). The symmetrical property of structural interactions constrains thus the motif analysis to consider a reduced number of motifs.

Whereas positive functional links are clearly correlated with the structural connections, results suggest that the local distribution of positive and negative links may play a role on the non-random organisation of brain networks. However, the precise relevance of such signed motifs on brain abilities remains a central question for future research. We therefore hope that our approach will foster more principled and successful analysis of multimodal brain connectivity datasets.

Chapter 4

Models of multiplex networks

In the previous chapters we characterised the structure of real-world multiplex networks by looking at the properties of nodes, edges, motifs and layers. How can we model and explain the emergence of some given features in our system? Historically, the investigation of the properties of networks from the real world has progressed together with increasing modelling efforts aiming at quantifying how special or peculiar were the observed patterns. For instance, how high clustering can coexist with a low diameter in a graph? Why are most networks characterised by a degree distribution with a fat tail? More in general, for a given feature how probable is to find them in an appropriately chosen family of random graphs, and which are the mechanisms that determine its appearance? In this Chapter we review a few classes of models of multiplex networks, namely the canonical and microcanonical ensembles, other static models aiming at reproducing real-world patterns of node activity, layer activity and edge overlap, and basic models of growing multiplex networks.

4.1 Canonical and microcanonical ensembles

A standard approach to study the structure of a given network is to quantify how probable is to observe a network with similar properties in an appropriately defined ensemble of random graphs whose elements satisfy certain constraints. For instance, it is a well-known fact that graphs with power-law degree distributions are extremely rare in the classical Erdös-Renyi random graph ensemble, where each pair of nodes are connected with a given probability p. As a consequence, the hypothesis that power-law degree distributions arise as a result of a uniform distribution of edges across the nodes can be safely rejected, and we can conclude that some other mechanism should be at work in

the formation of graphs with heterogeneous degree sequences.

An ensemble of graphs is determined by a set of constraints that its elements should satisfy. According to the type of constraints, we can identify at least two classes of random network ensembles, namely *canonical ensembles*, where each graph of the ensemble satisfies the set of constrains on average (soft constraints), and *microcanonical ensembles*, where each graph satisfies all the constraints exactly (hard constraints). It is possible to define a sequence of canonical and microcanonical ensembles of multiplex networks [14], where the constraints are just the average degree at each of the M layers, or the degree distribution of each layer, or the degree distribution together with the distribution of edge overlap, and so on.

Each multiplex networks ensemble is defined by providing the probability $P(\boldsymbol{A})$ for each of the possible configuration of multiplex networks $\boldsymbol{A}$ which satisfy the constraints. Starting from $P(\boldsymbol{A})$, the Shannon entropy of the ensemble is defined as

$$S = -\sum_{\mathcal{M}} P(\mathcal{M}) \ln P(\mathcal{M}) \tag{4.1}$$

[14].

For the special case of uncorrelated multiplex networks, we have

$$\langle a_{ij}^{[\alpha]} a_{ij}^{[\beta]} \rangle = \langle a_{ij}^{[\alpha]} \rangle \langle a_{ij}^{[\beta]} \rangle \tag{4.2}$$

the probability $P(\mathcal{M}) \equiv P(\boldsymbol{A})$ can be factorised into the probability of observing each single layer, i.e.

$$P(\boldsymbol{A}) = \prod_{\alpha=1}^{M} P^{[\alpha]}(A^{[\alpha]}). \tag{4.3}$$

In this particular case, the entropy of the multiplex ensemble reads

$$S = \sum_{\alpha=1}^{M} S^{[\alpha]} = -\sum_{\alpha=1}^{M} P^{[\alpha]}(A^{[\alpha]}) \ln(P^{[\alpha]}(A^{[\alpha]})). \tag{4.4}$$

In the following we focus on the canonical - indicated by C - and microcanonical - denoted by MC - ensembles of multiplex networks. Let us assume that we have T soft constraints such that

$$\sum_{\boldsymbol{A}} P(\mathcal{M}) F_{\mu}(\mathcal{M}) = C_{\mu} \tag{4.5}$$

where $\mu = 1, \ldots, T$, and $F_{\mu}(\mathcal{M})$ describes how such constraints are imposed on the network, such as the degree of each node of the network at each layer α, or the total number

of edges $K^{[\alpha]}$ for $\alpha = 1, \ldots, M$. The probability $P_C(\mathcal{M})$ of observing the multiplex $\mathcal{M}$ can be obtained by maximising the entropy S under the given set of constraints. By solving the optimisation problem one obtains:

$$P_C(\mathcal{M}) = \frac{1}{Z_C} \exp\left[- \sum_{\mu} \lambda_\mu \mathrm{F}_\mu(\mathcal{M}) \right] \tag{4.6}$$

where Z_C is the partition function of the canonical multiplex ensemble and the values of the Lagrangian multipliers λ_μ are obtained by satisfying Eq.4.5 imposing such functional form for $P_C(\mathcal{M})$. In the canonical multiplex ensembles we have:

$$S = \sum_{\mu} \lambda_\mu C_\mu + \ln Z_c. \tag{4.7}$$

Conversely, in the microcanonical multiplex ensemble each multiplex configuration compatible with the hard constraints has the same probability

$$P_{MC}(\mathcal{M}) = \frac{1}{Z_{MC}} \prod_{\mu=1}^{T} \delta[F_\mu(\mathcal{M}), C_\mu] \tag{4.8}$$

where δ is the Kronecker delta function and $Z_{MC} = \sum_{\boldsymbol{A}} \prod_{\mu=1}^{T} \delta[F_\mu(\mathcal{M}), C_\mu]$ is the microcanonical partition function, accounting for the number of multiplex networks satisfying the T hard constraints $F_\mu(\mathcal{M}) = C_\mu$. By defining the entropy of these ensembles as $N\Sigma$, such entropy reads

$$N\Sigma = \sum_{\boldsymbol{A}} P_{MC}(\mathcal{M}) \ln P_{MC}(\mathcal{M}) = \ln Z_{MC} \tag{4.9}$$

where Σ is the Gibbs entropy of the multiplex ensemble. It can be shown that the Gibbs entropy Σ is related to the corresponding Shannon entropy S by $N\Sigma = S - N\Omega$, where Ω is the logarithm of the probability that in the related canonical multiplex ensemble the hard constraints $F_\mu(\mathcal{M})$ are satisfied.

The author of Ref. [14] provided an exhaustive explanation of how the entropy and the partition function can be computed in different classes of multiplex networks with increasingly stringent sets of constraints, both for the canonical and for the microcanonical ensembles. The same approach has been generalised to a number of more complicated structures, including spatial multiplex networks [107] and multiplex networks with heterogeneous activities of the nodes [108]. The Authors of Ref. [109] study the canonical ensemble of the overlapping networks generated by merging different layers, where information on the connection between nodes is only accessible at the aggregated level.

4.2 A rewiring model for edge overlap

The canonical and microcanical ensembles do not account for edge correlations and the non-negligible values of edge overlap observed in real-world multiplex networks. Here we show how it is possible to construct a model for a 2-layer multiplex network where the degree distributions of each layer is kept fixed, and where one can control the edge overlap ω by rewiring a certain fraction r of the edges. The model was introduced in Ref. [110]. For simplicity, let us assume that the two layers have the same number of edges $K^{[\alpha]} = K^{[\beta]} = K$. If we start from two identical networks, we have maximum edge overlap $\omega = 1$. If we now keep fixed the structure on one of the two layers, and rewire one of the edges of the other layer, the number of links present in both layers decreases by one unit, while the number of those present in only one of the two layers increases by two units. Consequently, if we rewire a fraction r of the K edges of the second layer in such a way that each rewire decreases the number of edges existing on both layers, according to the definition of overlap provided in Eq. 2.10 we obtain:

$$\omega = \frac{(1-r)K}{(1+r)K} = \frac{(1-r)}{(1+r)} \tag{4.10}$$

By inverting such relation, we find that a given overlap ω corresponds to a rewire r equal to $r = (1-\omega)/(1+\omega)$. In practice, this model allows to obtain a prescribed value of edge overlap by rewiring a certain fraction r of the edges in one of the two layers.

4.3 Models of node and layer activity

The concept of node and layer activity is peculiar to multilayer networks, and it is interesting to assess whether simple models can give account for the observed heterogeneous distributions of node and layer activities. In the following we provide a brief review of some null models proposed so far to quantify the peculiarity of given distributions of node and layer activities.

Hypergeometric model (HM). Let us consider two layers α and β with $N^{[\alpha]}$ and $N^{[\beta]}$ active nodes respectively. If initially the two layers have no active nodes and we then sample uniformly at random from $\{1, 2, 3, \ldots, N\}$ $N^{[\alpha]}$ nodes on layer α and $N^{[\beta]}$ nodes on layer β and we activate them, then the probability that m of them are active at both layers follows a hypergeometric distribution

$$p(m; N, N^{[\alpha]}, N^{[\beta]}) = \frac{\binom{N^{[\alpha]}}{m}\binom{N-N^{[\alpha]}}{N^{[\beta]}-m}}{\binom{N}{N^{[\beta]}}}, \tag{4.11}$$

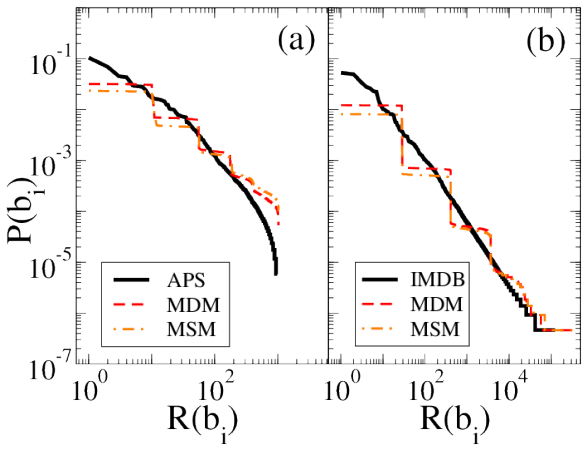


Figure 4.1: The rank distribution of node-activity vectors in APS (a) and IMDb (b), compared with those of synthetic multiplex networks genrated using MDM and MSM [22].

according to which the expected number of nodes active at both layers is equal to $N^{[\alpha]}N^{[\beta]}/N$, the expected pairwise multiplexity is

$$\tilde{Q}^{[\alpha,\beta]} = \frac{N^{[\alpha]}N^{[\beta]}}{N^2} \tag{4.12}$$

and the expected Hamming distance reads

$$\tilde{H}^{[\alpha,\beta]} = \frac{\sum_{m=0}^{N^{[\beta]}}(N^{[\alpha]} + N^{[\beta]} - 2m) \times p(m; N, N^{[\alpha]}, N^{[\beta]})}{\min(\mathrm{N}, \mathrm{N}^{[\alpha]} + \mathrm{N}^{[\beta]})} \tag{4.13}$$

This is the simplest model of node activation and is known as the hypergeometric model [22]. However, the authors of Ref. [22] have shown that the distribution of pairwise multiplexity and pairwise Hamming distance in real-world multiplex networks is not compatible with those given in Eq. 4.12 and Eq. 4.13.

Multi-activity Deterministic and Stochastic Models (MDM and MSM). Let us now consider the problem of constructing a multiplex networks with a fixed number of layers M, a fixed number of nodes N which are active on at least one of the M layers, and where each node i has an assigned node activity B_i, which is for instance set equal to the node activity observed in the real network. By sampling for each node one of the $\binom{M}{B_i}$ vectors of node-activity with B_i non-zero entries, the distribution of the total node activity of the original system is kept fixed, whereas the correlations in the layer activity and the distribution of the node-activity vectors are not preserved. Moreover, in such a

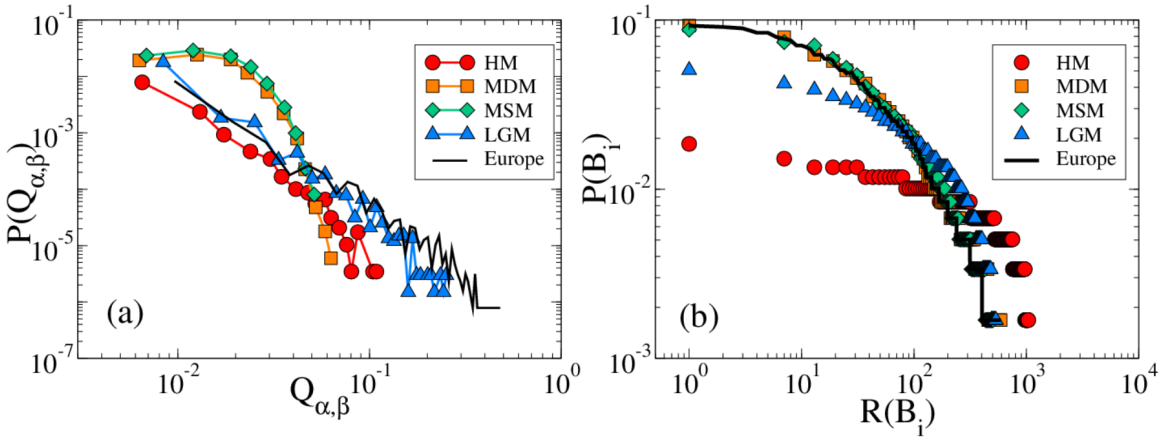


Figure 4.2: Distribution of pairwise multiplexity (a) and Zipf's plot of node-activity (b) for the European airlines multiplex network (solid black line) and the corresponding synthetic networks obtained by four different models, namely: HM (red circles), MDM (orange squares), MSM (green diamonds) and LGM (blue triangles). Notice that LGM fits well the distribution of pairwise multiplexity, and performs better than HM in reproducing the rank distribution of node-activity. The shape of $P(B_i)$ in MDM and MSM is identical to that of the original multiplex by construction [22].

model all layers have the same expected number of active nodes:

$$\tilde{N}^{[\alpha]} = \frac{1}{M} \sum_i B_i. \tag{4.14}$$

This model is known as the multi-activity deterministic model [22]. A variation of the model is constructed by activating node i in each layer α with probability $\bar{B}_i = B_i/M$, so that the expected activity of each layer stays the same but the original node-activity distribution is not preserved. This model is known as the multi-activity stochastic model [22].

A last model of layer activity is presented in the next section on growth models.

4.4 Growth models of multiplex networks

It is well-known that networks evolve and change over time. Historically, a lot of efforts have been devoted to the study of growing network models able to reproduce specific features. For instance, the scale-free degree distribution of many networks can simply emerge as a consequence of coupling together simple mechanisms such as *growth* and

preferential attachment [6]. In this section we review a few basic growth models for multiplex networks.

Layer Growth with Preferential Activation Model (LGM). The most simple example of this class is a model of layer-growth, aimed at explaining the fat-tail distribution of layer activity observed in empirical data [22]. The model works as follows. We start at time $t_0 = 0$ with a multiplex with M_0 layers and N nodes. At each time t, a new layer α joins the network with $N^{[\alpha]}$ nodes to be activated, where $N^{[\alpha]}$ can be observed from the data-set we are attempting to reproduce. Each node i has then a probability to be active on that layer at time t equal to:

$$p_i(t) = A + B_i(t), \tag{4.15}$$

where $B_i(t)$ is the number of layers where node i is already active and $A > 0$ is a constant that allows the activation of nodes not yet active in the multiplex. When the number of layers in the model increases, the distribution of layer activity $P(N^{[\alpha]})$ approaches a power law.

Linear and Non-linear preferential attachment. Another important class of growth models is that where not layers, but individual nodes join sequentially the network, for instance by connecting to preexisting vertices on possibly different layers. In such regard, it is clear that the specific shape of the attachment function determines the long-term statistical properties of the final multiplex graph. In single-layer networks, a particularly well-studied case is the so-called preferential attachment, where nodes choose to attach to older vertices depending on a function (in the simplest case linear) of their degree k [6]. In a multiplex network the degree of each node j is a vector and the probability $\Pi_{i\to j}^{[\alpha]}$ that a new node i attaches to j on a given layer α in general depends on all its components. In formula:

$$\Pi_{i\to j}^{[\alpha]} = \frac{F_j^{[\alpha]}(\mathbf{k}_j)}{\sum_\ell F_\ell^{[\alpha]}(\mathbf{k}_\ell)} \tag{4.16}$$

The most simple class of preferential attachment models is obtained by considering linear attachment kernels, i.e. by setting $F_j^{[\alpha]}$ as a convex combination of the degrees of node j at all the layers [111, 112]. The interesting result is that linear attachment kernels produce multiplex networks whose layers have power-law degree distributions, but where inter-layer degree correlations are always positive, meaning that a hub on one layer is also a hub on the other layer as well. This is due to the fact that the expected final degree of a node on a certain layer is determined solely by the time at which it

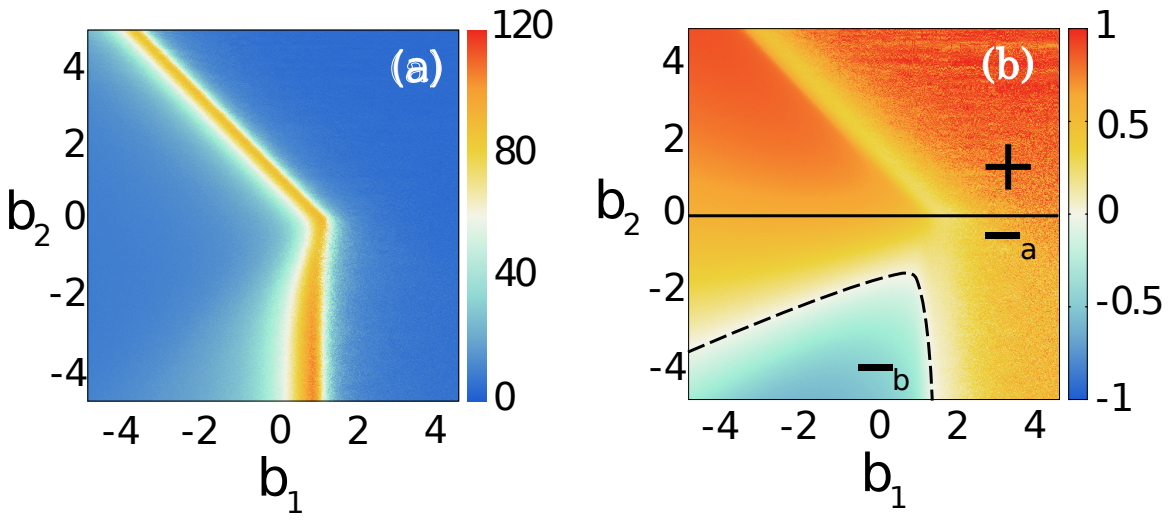


Figure 4.3: As a function of the two parameters b_1 and b_2 we report, by means of a color code (a) the number of distinct degree classes $|k|$ (a) the Kendall's τ correlation coefficient (b). The solid black lines in panel (a) separate the non-condensed (region I) from the condensed phase (region II, small $|k|$). The solid black line in panel (b) separates the two regions with positive (region $+$) and negative inter-layer degree correlations (region $-$), respectively corresponding to $b_2 > 0$ and $b_2 < 0$. The value of τ for the whole multiplex is negative only in region $-_{\boldsymbol{b}}$ [115].

joins the network [111]. A generalisation of the closed-form solutions for the joint degree distribution of heterogeneously growing multiplex networks with arbitrary number of layers and arbitrary times can be found in [113].

A more interesting class of multiplex networks is obtained by considering non-linear attachment kernels [114]. The authors of Ref. [115] started from the case of multiplex networks with two layers, using the attachment kernel:

$$F_j^{[1]} \propto \left(k_j^{[1]}\right)^{b_1} \left(k_j^{[2]}\right)^{b_2} \tag{4.17}$$

where $b_1, b_2 \in \mathbb{R}$. By tuning the relative values of the exponents b_1 and b_2, one can obtain multiplex networks where each layer has either an exponential, a power-law, or a condensed degree distribution (where super-hubs with extensive degrees appear). Moreover, with non-linearity it is possible to get non only positive inter-layer degree correlations as in the case of the linear model, but also null and negative correlations that have been observed in real-world systems [22]. In the same work the authors suggested several possible generalisations of the model to the case of multiplex networks with $M \leq 2$ layers. An interesting model of multiplex network growth which takes into account

weighted links, aimed at reproducing the structure of some layered social networks, can be found in Ref. [116].

4.5 Discussion

In this chapter we reviewed a number of models of multiplex networks. We first considered ensembles of graphs which satisfy on average or exactly several constraints, such as the number of links $K^{[\alpha]}$ at each layer α, and which are known respectively as the canonical and the microcanonical ensembles. A second set of static models was suggested to reproduce some specific features such as node and layer activity, as well as edge overlap. An important class of models is that of growth multiplex networks, where nodes join the system at different time. In particular, we focus on generalisation to the multiplex case of models based on linear and non-linear preferential attachment, where new nodes attach to vertices already in the system with a probability which depends on a function of their degree. Simple preferential attachment models, while able to reproduce some empirical patterns such as inter-layer degree correlations, do not allow to construct multiplex networks with community structure. In the next Chapter we first evaluate community structure into and across the layers of real collaboration networks, and then suggest a more sophisticated - but still simple - growth model, which based on intra-layer and inter-layer triadic closure mechanisms is able to produce multiplex networks with tuneable community structure [63].

Chapter 5

Community structure

More often than not the agents of a social system prefer to combine their efforts in order to achieve results that would be otherwise unattainable by a single agent alone. A relevant role in the organisation of such systems is therefore played by the emerging patterns of collaboration within a group of individuals, which have been widely and thoroughly investigated in the last few decades [48, 117]. In a collaboration network, two individuals are considered to be linked if they are bound by some form of partnership. For instance, in the case of scientific collaborations, the nodes of the networks correspond to scientists and the relationship between two authors is testified by the fact that they have co-authored one or more papers [43]. Another well-known example of collaboration network is that of co-starring graphs, where the nodes represent actors and there is a link between two actors if they have appeared in the same movie.

The study of large collaboration systems has revealed the presence of a surprisingly high number of triangles in the corresponding networks [5, 118]. This indicates that two nodes with a common neighbour have a higher probability to be linked than two randomly chosen nodes. This effect, known as *transitivity* [48], can be easily explained in terms of a basic mechanism commonly referred to as *triadic closure* [119], according to which two individuals of a collaboration network have a high probability to connect after having been introduced to each other by a mutual acquaintance [5, 120, 121]. Some other works have pointed out that triadic closure can also explain other empirical features of real-world collaboration networks, including fat-tailed degree distributions and correlations between the degrees of neighbouring nodes [122, 123].

Another remarkable feature often observed in social and collaboration networks is the presence of meso-scale structures in the form of *communities*, i.e. groups of tightly connected nodes which are loosely linked to each other [124]. Interestingly, structural

communities quite often correspond to functional groups [125].

An important observation is that not all the links of a collaboration network are equal, since collaborations can often be classified into a number of different categories. For instance, scientific co-authorship can be classified according to the research field, while actors often appear in movies of different genres. In these cases, a collaboration network is indeed better described in terms of multiplex network. In the previous chapter, we reviewed several static and growing models of multiplex networks, such as multiplex network ensembles, models of node and layer activity and growing models based on linear [111, 112] or non-linear [115] preferential attachment, or on weighted networks [116]. Less attention has so far been devoted to define and extract communities in multiplex networks [126, 127].

In this chapter we investigate the multiplex nature of communities in collaboration networks and we propose a simple model to explain the appearance, coexistence and co-evolution of communities at the different layers of a multiplex [63]. Our hypothesis is that the formation of communities in collaboration networks is an intrinsically multiplex process, which is the result of the interplay between intra-layer and inter-layer triadic closure. For instance, in the case of scientific collaborations, multiplex communities naturally arise from the fact that scientists may collaborate with other researchers in their principal field of investigation and with colleagues coming from other scientific disciplines. Analogously, actors can prefer either to specialise in a specific genre or instead to explore different (sometimes dissonant) genres, and these two opposite behaviours undoubtedly have an impact on the kind of meso-scale structures observed on each of the layers of of the system. The generative model we propose here mimics two of the most basic processes that drive the evolution of collaborations in the real world, namely intra- and inter-layer triadic closure, and is able to explain the appearance of overlapping modular organisations in multi-layer systems. We will show that the model is able to reproduce the salient micro-, meso- and macro-scale structure of different real-world collaboration networks, including the multi-layer network of co-authorship in journals of the American Physical Society (APS) and the multiplex co-starring graph obtained from the Internet Movie Database (IMDb).

5.1 Multiplex communities in collaboration networks

We start by analysing the structure of two multiplex collaboration networks from the real world. The first multiplex is constructed as a subset of the APS co-authorship data set, first present in Chapter 1 and consists of four layers representing four sub-

fields of physics (respectively, Nuclear physics, Particle physics, Condensed Matter I, and Interdisciplinary physics). In particular, we considered only scientists with at least one publication in each of the four sub-fields, and we connected two scientists at a certain layer if they had co-authored at least a paper in the corresponding sub-field. The second multiplex is constructed as a subset of the Internet Movie Database (IMDb), again presented in Chapter 1, and consists of four layers respectively representing the co-starring networks of actors with at least one participation in four different genres, namely Action, Crime, Romance, and Thriller movies. The basic structural properties of each layer of the two 4-layer multiplexes are summarised in Table 5-A.

APS	N	$\langle k \rangle$	C
Nuclear (N)	1238	4.75	0.27
Particle (P)	1238	4.66	0.30
Cond. Matt. I (CM)	1238	10.29	0.24
Interdisciplinary (I)	1238	7.37	0.26
IMDb	N	$\langle k \rangle$	C
Action (A)	55797	83.56	0.61
Crime (C)	55797	82.30	0.58
Romance (R)	55797	86.00	0.59
Thriller (T)	55797	77.75	0.56

Table 5-A: We report the number of nodes N, the average degree $\langle k \rangle$ and the clustering coefficient C for each layer of a subset of the APS and IMDb data sets. In particular, we focus on the multiplex collaboration network of all scientists active in Nuclear, Particle, Condensed Matter I and Interdisciplinary physics, and the multiplex collaboration network of all actors starring in Action, Crime, Romance and Thriller movies. All the layers of APS have a clustering coefficient C in the range $[0.24, 0.30]$. Conversely, the values of C of all the IMDb layers are in the range $[0.56, 0.61]$ [63].

Since we are interested in assessing the role of intra- and inter-layer triadic closure in the formation of meso-scale multiplex structures, we quantified the transitivity of each layer through the clustering coefficient C [5]. We notice that the four layers of each data set have similar values of clustering, ranging respectively in $[0.24, 0.3]$ in the case of APS and in $[0.56, 0.61]$ for IMDb. As we will discuss in the following, by focusing on layers having comparable clustering we will be able to perform a comparison between the structure of these real-world multiplex networks and the proposed model in its simplest formulation.

The multiplex nature of communities in collaboration networks can be measured by means of the normalised mutual information (NMI) [128], which quantifies the similarity between the partition in communities observed in two different layers of a multiplex. In

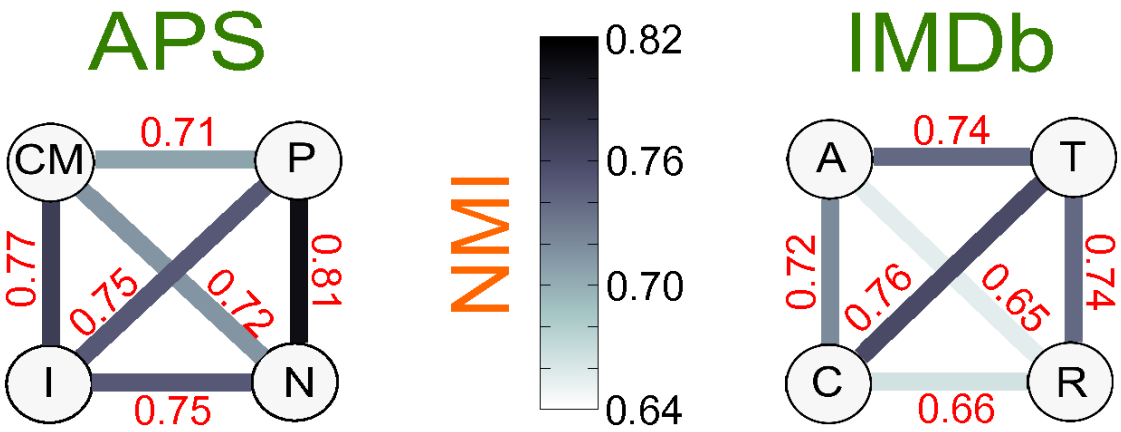


Figure 5.1: We report the similarity among communities at the different layers of two real-world collaboration networks.In each of the two graphs nodes represent the layers of the multiplex (APS on the left and IMDb on the right) and the edges are coloured according to the value of the normalised mutual information for the community decompositions at the corresponding pairs of layers [63].

particular, given the two partitions $\mathcal{P}_\alpha$ and $\mathcal{P}_\beta$ respectively associated to layer α and layer β, we denote the normalised mutual information (NMI) between them as

$$NMI(\mathcal{P}_\alpha, \mathcal{P}_\beta) = \frac{-2\sum_{m=1}^{M_\alpha}\sum_{m'=1}^{M_\beta} N_{mm'} \log\left(\frac{N_{mm'}N}{N_m N_{m'}}\right)}{\sum_{m=1}^{M_\alpha} N_m \log\left(\frac{N_m}{N}\right) + \sum_{m'=1}^{M_\beta} N_{m'} \log\left(\frac{N_{m'}}{N}\right)} \tag{5.1}$$

where $N_{mm'}$ is the number of nodes in common between module m of partition $\mathcal{P}_\alpha$ and module m' of partition $\mathcal{P}_\beta$, while N_m and $N_{m'}$ are respectively the number nodes in module m and in module m'. The partition in communities on each layer has been obtained through the algorithm Infomap [129]. The normalised mutual information takes values in $[0, 1]$. In general, higher values of NMI correspond to more similar partitions. The values of NMI for each pair of layers in APS and IMDb are shown in Fig. 5.1. It is interesting to notice that in general pairs of layers corresponding to related subjects or genres exhibit higher values of NMI. This is for instance the case of Nuclear Physics and Particle Physics in APS. Similarly, in the IMDb network we observe a higher similarity between the communities at the three layers representing respectively Thriller, Crime and Action genres. Conversely, the layer of Romance movies displays a different modular structure from Crime and Action. Notice also that the level of similarity between the communities of two layers can vary substantially, despite the four layers of each multiplex have roughly the same clustering coefficient.

5.2 Growth model with community structure

In this Section we introduce a model to grow collaboration networks with tunable multiplex community structure, able to reproduce the patterns observed in the considered real-world systems [63]. Let us consider for simplicity the case of a multiplex with $M = 2$ layers, and assume that initially each layer consists of a clique of n_0 nodes. Then at each time step t a new node is added to the network, with $m^{[1]}$ edge stubs to be connected on layer 1 and $m^{[2]}$ other stubs to be connected on layer 2. The multiplex network grows according to the following rules:

- *Layer selection.* The newly arrived node i selects one of the two layers $\{1, 2\}$ uniformly at random. Let us label the first selected layer with the index a. The first edge of i is connected to one of the existing nodes on that layer, chosen uniformly at random, that we call n_a.

- *Intra-layer triadic closure (I).* The remaining $m^{[a]}$-1 edges of node i on layer a are attached with probability $p^{[a]}$ to one of the first neighbours of n_a, chosen uniformly at random, and with probability $1 - p^{[a]}$ to one of the nodes of layer a, chosen uniformly at random.

- *Inter-layer triadic closure.* When all its $m^{[a]}$ edges on layer a have been created, node i starts connecting on the other layer b with $m^{[b]}$ edges. The first link in layer b is created with probability p^* to the same node n_a, and with probability $1 - p^*$ to one of the other nodes, chosen uniformly at random. The node to which this first link is attached is called n_b.

- *Intra-layer triadic closure (II).* The remaining $m^{[b]}$-1 links at layer b are attached with probability $p^{[b]}$ to one of the first neighbours of n_b chosen uniformly at random, and with probability $1 - p^{[b]}$ to one of the nodes at layer b, chosen uniformly at random.

This general model has five parameters to be tuned, namely the number of new edges $m^{[1]}$ and $m^{[2]}$ brought by each new node on each of the two layers, which determine the average degree on each layer, and the three probabilities $p^{[1]}$, $p^{[2]}$, and p^*, which are respectively responsible for the formation of intra- and inter-layer triangles. In fact, by varying the parameters $p^{[1]}$ and $p^{[2]}$ we can tune the strength of the intra-layer triadic closure mechanism, i.e the probability to form triangles on each of the two layers. In particular, larger values of $p^{[1]}$ and $p^{[2]}$ will foster the creation of a larger number of triangles in layer 1 and layer 2 respectively. Conversely, the parameter p^* tunes the inter-layer triadic closure mechanism, and in particular high values of p^* correspond to

a higher probability that the neighbourhoods of node i at the two layers will exhibit a certain level of overlap. These two simple attachment rules, namely intra-layer and inter-layer triadic closure, aim to describe the real mechanisms characterising the evolution of collaboration networks. We argue that, for instance, scientists do not tend to collaborate with other scientists at random. Instead, they usually exploit the neighbourhoods of their collaborators in a specific field (*intra-layer* triadic closure). Similarly, when opening themselves to new scientific fields, a researcher usually takes into account the neighbourhoods of their past colleagues from previous collaborations in other fields (*inter-layer* triadic closure). A schematic representation of the model is depicted in Fig. 5.2.

It has been recently shown [123] that in a single-layer network scenario the interplay between random attachment and triadic closure leads to a network growth in which the attachment probability (i.e., the probability for an existing node to receive one of the new edges) is a sub-linear function of the degree, and produces networks with non-trivial community structure, as long as the link density is not too high. In the multi-layer model we propose, the further addition of an inter-layer triadic closure mechanism allows to tune at will the overlap between the community structures at the different layers.

5.3 Validation in a simple scenario

To assess the ability of the model to reproduce the organisation of communities in multiplex networks, we start by considering a simple scenario, i.e. the case in which the layers of the multiplex have the same density ($m^{[1]} = m^{[2]} = m$) and the same clustering coefficient ($p^{[1]} = p^{[2]} = p$). We show that this simplified version of the model is already able to reproduce both the different levels of similarity between community structures at different layers, and the microscopic patterns of intra-layer and inter-layer degree correlations observed in the real-world collaboration multiplexes of APS and IMDb [63].

In Fig. 5.3(a), we report the values of the clustering coefficient C (which, by construction, does not depend on the parameter p^*) for several realisations of the model. In particular, we run our model to build networks with $N = 20000$ nodes, for each pair of values (p, p^*) As expected, the clustering coefficient of each layer is a linearly increasing function of the parameter p, which tunes the strength of intra-layer triadic closure. This means that, if we consider a real-world multiplex network whose layers have approximately the same value of clustering coefficient C, we can set the value of the parameter p of the model accordingly. This is for instance the case of the four-layer multiplex networks of APS and IMDb constructed in the previous Section, where all the

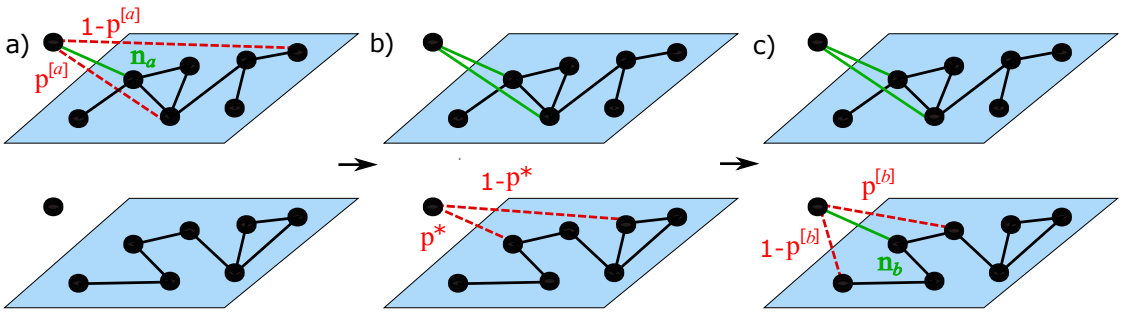


Figure 5.2: We show a schematic representation of network growth with intra-layer and inter-layer triadic closure. A newly arrived node i creates $m^{[1]}$ new edges on layer 1 and $m^{[2]}$ new edges on layer 2. The new node starts by choosing at random one of the two layers $\{1, 2\}$. We indicate the first chosen layer using the label a. a) The first link of the new node is connected to one of the nodes of layer a, chosen uniformly at random and called n_a (solid green line). Each of the remaining $m^{[a]} - 1$ links is attached with probability $p^{[a]}$ to a neighbour of the previously chosen node (intra-layer triadic closure) or with probability $1 - p^{[a]}$ to one of the nodes at layer a, chosen uniformly at random (dashed red lines). b) Afterwards, the new node starts connecting on the other layer b. The first link on layer b is created to node n_a with probability p^*, or to one of the other nodes at layer at random with probability $1 - p^*$. We call n_b the first node to which i attaches on layer b. c) Each of the $m^{[b]} - 1$ remaining edges on layer b are attached with probability $p^{[b]}$ to one of the neighbours of n_b, and with probability $1 - p^{[b]}$ to one of the nodes on layer b, chosen uniformly at random [63].

layers have comparable levels of clustering. We obtain $p = 0.40$ for APS and $p = 0.85$ for IMDb, respectively.

In Fig. 5.3(c) we show, as a colour-map, the values of NMI of the networks obtained through the proposed model by using different combinations of the parameters p and p^*. It is evident that, in spite of its simplicity, the model can yield a quite rich variety of multiplex networks. In agreement with intuition, when both p and p^* are large one obtains multiplexes with higher values of NMI. In fact, in this regime both the intra-layer and inter-layer triadic closure mechanisms are strongly affecting the network evolution and, as a consequence, it is likely that the new node joining the network will close a triad on both layers in the same region of the network. As a consequence, each layer will have a strong community structure (large p) which is pretty much correlated to the one present on the other layer, due to the large value of inter-layer triadic closure p^*. Conversely, if the inter-layer parameter p^* is small we will obtain layers whose partitions in communities are poorly correlated when p is large (blue region in the phase space

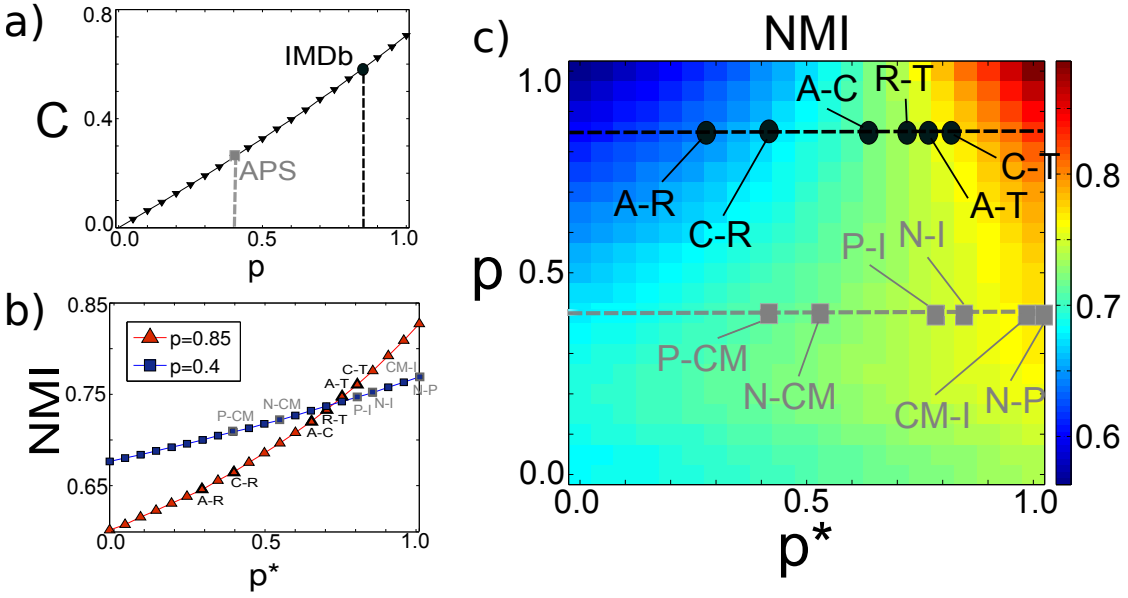


Figure 5.3: We report here the model calibration in the simplest scenario and the values of p and p^* extracted for the different pairs of layers of the four-layer collaboration networks of APS and IMDb. (a) The clustering coefficient C depends exclusively on the parameter p, which tunes intra-layer triadic closure. Since all the layers of those two multiplex networks have comparable clustering coefficients, we are able to determine the value of the parameter p in each of the two cases. (b) For each pair of layers, we can also determine the value of the inter-layer triadic closure parameter p^* by setting it equal to the value which yields an organisation in communities characterised by a value of NMI compatible with that observed in the real network [63].

of Fig. 5.3, while the NMI is only marginally larger when p is very small (bottom-left corner of the phase space).

In Fig. 5.4 we report two realisations of the multiplex network model with $N = 50$, $m^{[1]} = m^{[2]} = 2$ and $p^{[1]} = p^{[2]} = 0.9$, respectively for $p^* = 0.9$ (left) and $p^* = 0.1$ (right). Nodes belonging to the same community are reported using the same colour, and the colour chosen for each community in the second layer (bottom) corresponds to the colour of the community in the first layer (top) for which the node overlap between the communities is maximum. These two examples help explain the role of the parameter p^* in shaping the inter-layer modular structure of the network. For $p^* = 0.9$ (left panel) the community structures of the two layers are closely matched (this situation corresponds to the high values of NMI found in the top-right region of the heat-map in Fig. 5.3), while for $p^* = 0.1$ (right panel) the communities at the two layers are uncorrelated (low values of NMI in the top-left of the heat-map in Fig. 5.3).

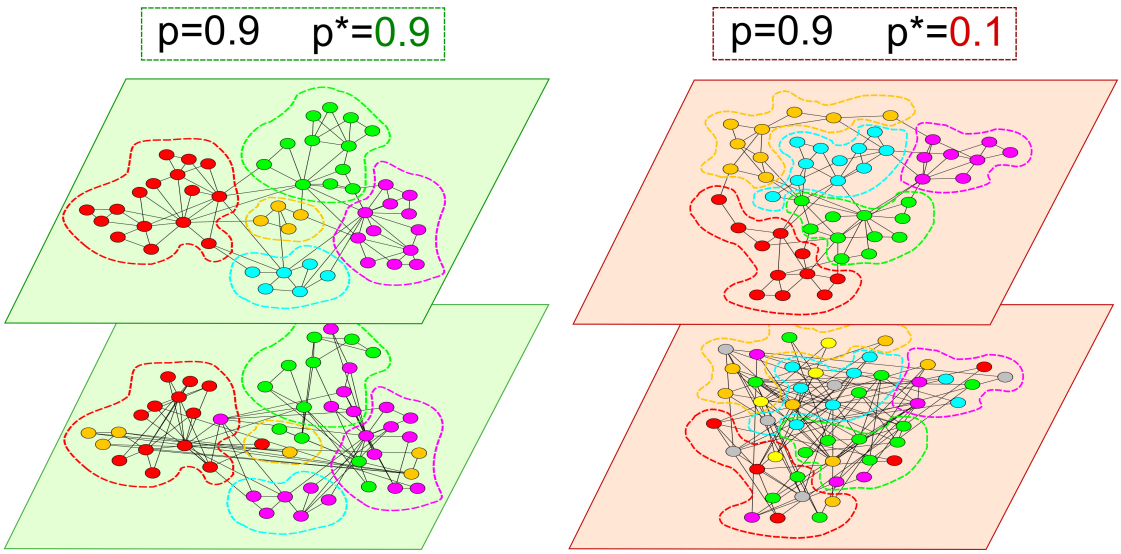


Figure 5.4: We show the effect of the value of the inter-layer triadic closure parameter p^* on the multiplex community structure. The two top layers show two typical realisations of the simplest version of the network model with $N = 50$, $m^{[1]} = m^{[2]} = 2$ and $p^{[1]} = p^{[2]} = 0.9$. Nodes belonging to the same community are given the same colour and are drawn close to each other. The two layers at the bottom of each multiplex are obtained by setting, respectively, $p^* = 0.9$ (left) and $p^* = 0.1$ (right). The nodes maintain the same placement in space on the second layer, but are coloured according to the community they belong in that layer (colours are chosen in order to maximise the number of nodes that have the same colour in the two layers). It is evident that the community structures of the two layers on the left, corresponding to $p^* = 0.9$, are very similar, while the partition into communities of the upper layer on the left panel is substantially different from the one observed in the bottom layer of that multiplex [63].

Differently from the clustering coefficient C, the values of the normalised mutual information NMI depend on both p and p^*. Having already determined a candidate value of p for each multiplex by fitting the clustering coefficient of its layers, we can determine the strength of the inter-layer triadic closure mechanism by fitting the NMI. Remarkably, for any fixed value of p, the simplest formulation of our model is able to reproduce all the values of NMI observed in the real-world networks by just tuning the parameter p^*, with the exception of the pair Nuclear-Particle physics which is slightly out of the plane with an NMI value of 0.81 (represented on the right border of the plane which corresponds to NMI=0.79). We would like to note here that the model is able to produce a remarkably wide range of values of NMI, which span the whole interval $[0.6, 0.9]$.

5.4 Calibration for generic multiplex networks

We now discuss how to calibrate the model in the most general case in which the layers might possibly have different edge density, i.e. $m^{[1]} \neq m^{[2]}$, and different clustering, i.e. $p^{[1]} \neq p^{[2]}$ [63]. As an example, we consider the co-authorship networks of the same four sub-fields of physics (namely, Nuclear, Particle, Condensed Matter I and Interdisciplinary physics) used to construct the four-layer APS multiplex (cf. Table 5-A and Fig 5.1). However, we focus here on all two-layer multiplex networks obtained by combining two networks at a time, so that, for instance, a node appears in the Nuclear-Particle (N-P) multiplex network if the corresponding author has published papers in both sub-fields. In general, the obtained multiplex networks are composed by layers with different edge density and different clustering coefficients, as shown in Table 5-B, thus we need to set separately the four parameters of the model $p^{[1]}$, $p^{[2]}$, $m^{[1]}$ and $m^{[2]}$.

We start by observing that the average degree of a synthetic layer is $\langle k \rangle \simeq 2m$, where m is the number of edge stubs connected by a newly arrived node, so that the parameters $m^{[1]}$, $m^{[2]}$ of the model can be set respectively equal to $\left[\frac{\langle k^{[1]} \rangle}{2}\right]$ and $\left[\frac{\langle k^{[2]} \rangle}{2}\right]$, where $\langle k^{[1]} \rangle$ and $\langle k^{[2]} \rangle$ are the measured average degrees of the two layers (numbers are approximated to the closest integers). Similarly, as we show in Fig. 5.5(a), the clustering coefficient $C^{[\alpha]}$ of a layer α is univocally determined by $p^{[\alpha]}$, as soon as $m^{[\alpha]}$ is fixed. In Fig. 5.5(a) we show how the values of $C^{[\alpha]}$ change as a function of $p^{[\alpha]}$, for different values of $m^{[\alpha]}$. Hence, the values of the intra-layer triadic closure parameters $p^{[1]}$ and $p^{[2]}$ can be set in order to match the values of clustering coefficient observed in each of the two layers. The only parameter yet to be determined is p^*. However, if we set the values of $m^{[1]}$, $m^{[2]}$, $p^{[1]}$, and $p^{[2]}$ to match the densities and clustering coefficients of the layers, we can then run the model for different values of p^* and look for the one which yields a value of NMI as close as possible to the one observed in the real two-layer multiplex. This procedure is sketched in Fig. 5.5 (b) for the six two-layer multiplexes in APS.

In order to better understand the role of the different parameters, in Fig. 5.5(c) we report the values of NMI obtained from different realisations of the model with $m^{[1]} = m^{[2]} = m$ and $p^{[1]} = p^{[2]} = p$ for m varying in $[2, 3, ..., 10]$, and p varying in $[0, 0.1, ..., 1]$ at different values of p^*, $[0.05, 0.5, 0.95]$, corresponding respectively to low, intermediate and high inter-layer triadic closure strength. We see that the effect of the increase in the link density m of the layers leads to a decrease in the similarity of their community structures even for high values of p and p^*.

It is interesting to notice that, although the generic version of the model depends on five parameters, respectively accounting for layer density ($m^{[1]}$ and $m^{[2]}$), triadic closure

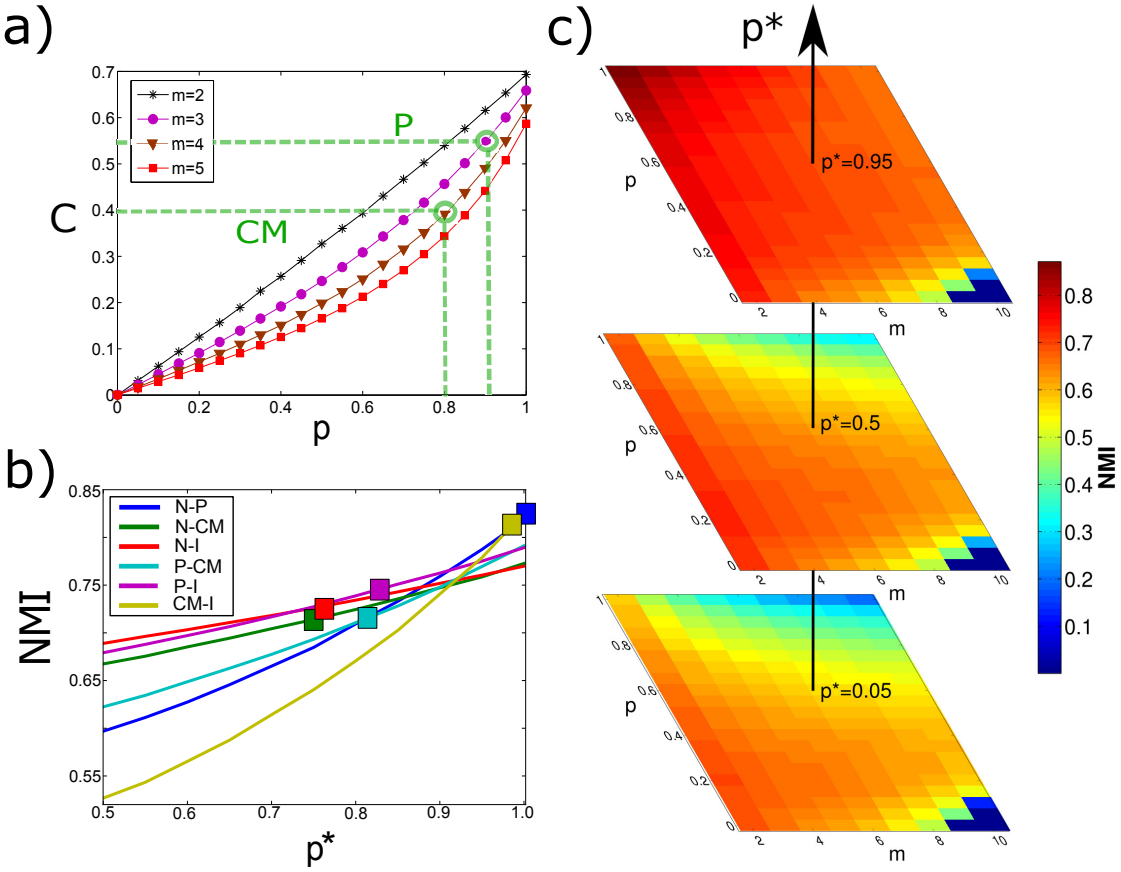


Figure 5.5: We show here the model calibration for generic multiplex networks. In panel a) we show the dependence of the clustering coefficient C on the intra-layer triadic closure parameter p for different values of the parameter m, which sets the layer's average degree. In the multiplex consisting of the layers Particle (P) and Condensed Matter I (CM), the average degree of each layer corresponds, respectively, to $m^{[1]} = 3$ and $m^{[2]} = 4$. The value of $p^{[1]}$ and $p^{[2]}$ are determined to match the clustering coefficients $C^{[1]}$ and $C^{[2]}$. In panel b), after having determined $m^{[1]}$, $m^{[2]}$, $p^{[1]}$ and $p^{[2]}$ for all the pairs of layers in the APS dataset, we run the model with such parameters for different value of p^* and infer, for each pair, the value of the inter-layer triadic closure parameter p^* yielding a value of NMI compatible with that observed (see Table 5-A for layers' acronyms). In panel c) we plot a heat-map of the NMI as a function of p and m, respectively for low (0.05), intermediate (0.50) and high (0.95) values of p^* in the model with $m^{[1]} = m^{[2]} = m$ and $p^{[1]} = p^{[2]} = p$. An increase in the link density of the layers produces a less correlated community structure in the two layers, even if the inter- and intra-layer triadic closure strengths are high [63].

($p^{[1]}$ and $p^{[2]}$), and inter-layer overlap of communities (p^*), the values of those parameters can be easily set by measuring just the average degree and the average clustering coefficient of each layer, and the normalised mutual information between the community structures at the two layers. Again, the good agreement between the synthetic networks and the real-world datasets extends also to other structural properties, such as intra-layer and inter-layer degree correlations. We remark that it is known that on each layer a minimum level of degree correlations is needed to achieve a certain level of clustering, as discussed in Ref. [130]. These results suggest that triadic closure plays an unexpectedly central role in determining the structural properties of real-world multiplex collaboration networks.

Layer 1	**Layer 2**	N	$\langle k^{[1]} \rangle$	$\langle k^{[2]} \rangle$	$C^{[1]}$	$C^{[2]}$	NMI
Nuclear	Particle	6572	6.88	7.46	0.56	0.56	0.83
Nuclear	Cond. Matt. I	3828	4.53	7.20	0.43	0.34	0.71
Nuclear	Interdisciplinary	2556	4.15	5.39	0.37	0.33	0.72
Particle	Cond. Matt. I	3774	5.70	7.82	0.53	0.40	0.71
Particle	Interdisciplinary	2502	4.82	5.66	0.49	0.39	0.74
Cond. Matt. I	Interdisciplinary	27257	10.34	7.05	0.55	0.64	0.82

Table 5-B: **Basic properties of duplex networks in APS.** We consider all the possible multiplex networks with $M = 2$ layers obtained from combinations of the APS collaboration networks corresponding to the four sub-fields Nuclear, Particle, Condensed Matter I and Interdisciplinary Physics. For each duplex, we report the number of nodes N, the average degree on the two layers $\langle k^{[1]} \rangle$ and $\langle k^{[2]} \rangle$, and the values of the clustering coefficients $C^{[1]}$ and $C^{[2]}$ [63].

5.5 Discussion

Human collaboration patterns are inherently multifaceted and often consist of different interaction layers. Scientific collaboration is probably the most emblematic example. As a Ph.D. student you usually join the scientific collaboration network by publishing the first paper with your supervisor in a specific field. Afterwards, you start being introduced by your supervisor to other researchers in the same field, e.g. to some of his/her past collaborators, and you might end up working with them, creating new triangles in the collaboration network of your field (what we called intra-layer triadic closure). But it is also quite probable that some of your past collaborators will in turn introduce you to researchers working in another -possibly related- area (what we called an inter-layer triadic closure), so that you will easily find yourself participating in more than just one field, and the collaboration network around you will become multi-dimensional. Such

multi-level collaboration patterns appear not to be specific of scientific production only, but are instead found in many aspects of human activity.

The multi-layer network framework provides a natural way of modelling and characterising multidimensional collaboration patterns in a comprehensive manner. In particular, we have argued that one of the classical mechanisms responsible for the creation of triangles of acquaintances, i.e. triadic closure, is indeed general enough to give also account for another interesting aspect of multi-level collaboration networks, namely the formation of cohesive communities spanning more than a single layer of interaction. It is quite intriguing that the simple model we proposed in this work, based just on the interplay between intra- and inter-layer triadic closure, is actually able to explain much of the complexity observed in the micro- meso- and macroscopic structure of multidimensional collaboration networks of different fields (science and movies), including not just transitivity but also intra- and inter-layer degree correlation patterns and the correspondence between the community partitions at difference layers. We also remark that such levels of accuracy in reproducing the features of real-world systems have been obtained without the introduction of ad-hoc ingredients.

The results reported in this chapter suggest that, despite the apparent differences in the overall dynamics driving scientific cooperation and movie co-starring, triadic closure is a quite generic mechanism and might indeed be one of the fundamental processes shaping the structure of multi-layer collaboration systems [63]. These findings fill a gap in the literature about modelling growing multidimensional networks, and pave the way to the exploration of other simple models which can help underpinning the driving mechanisms responsible for the emergence of complex multi-dimensional structures.

Chapter 6

Core-periphery structure

In complex network theory a *core* consists of a group of central and densely connected nodes which often control the overall behaviour of the system. Since the seminal paper by Borgatti and Everett [131], the core-periphery structure is recognised as one of the key meso-scale structures in complex networks, and was shown to be present in several real-world systems, such as the world trade web [132], many social [133] and biological networks [134], and the brain [135]. In this Chapter we explore the existence and the overlap between the cores of multiple interaction layers in a multilayer network, and propose a simple method to partition the nodes into a multiplex core-periphery structure by taking into account at the same time the connectivity of all the different layers without neglecting their specificity [**?**].

6.1 Block structure

The best way to visualise the core of a network is to consider its adjacency matrix. In a core-periphery structure, each node is assigned to one out of two possible categories, namely the core C or the periphery P. Nodes of type C are supposed to have a high chance to be connected among each other, as they form the "gravity centre" of the network. Conversely nodes of type P are supposed to have few connections, and in particular very low chances to be connected among each other.

An idealised representation of a core periphery structure in a network is shown through the adjacency matrix in Fig. 6.1(a), as first presented in Ref. [131]. We clearly distinguish the existence of different blocks. A first block (that we name 1) at the top left encodes information among connections among all nodes of type C. This group of

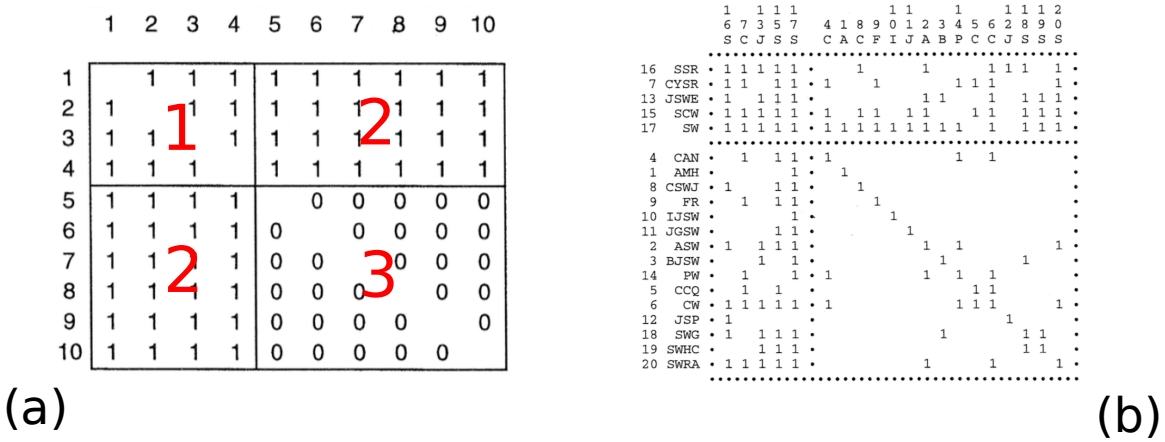


Figure 6.1: (a) Idealised core-periphery structure in networks: two core nodes are linked with probability $p_1 = 1$, a core node and a peripheral node are linked with probability $p_2 = 1$, two peripheral nodes are linked with probability $p_3 = 0$. (b) Core-periphery structure in a co-citations network among social work journals over a 1 year period across 1985-1986, as shown in Ref. [131].

nodes forms a fully connected clique such that this part of the adjacency matrix is filled exclusively with ones. Conversely, at the bottom right we observe a block (that we name 3) filled only with zeros, which represents the (absence of) connections among nodes of type P. By assuming we are dealing with an undirected network, the bottom left and the top right blocks encodes connections between nodes of type P and node of type C (these two symmetric sections of the adjacency matrix are named block 2). Also block 2 is filled exclusively with ones. As it is clear from this simplified picture, in the approach followed by Borgatti and Everett a node belonging to the periphery of the network is not necessarily a poorly connected node, but its connections are supposed to be towards core nodes only. Of course this representation is idealised and real-world systems can rarely be partitioned in a way to recover such block structure in the adjacency matrix. A core is almost never formed by a fully connected clique, and few connections among peripheral nodes may exist. Still, this simple block model was suggested as a reference to evaluate how good is a given partition of the nodes of the system in terms of a core-periphery structure. Formally, the Authors of Ref. [131] introduce the quality function

$$\Gamma = \sum_{i,j}^{N} a_{ij} C_{ij}, \tag{6.1}$$

where $C_{ij} = 1$ if both i and j are nodes of type C. Such number is compared with what is obtained in a network where the same number of nodes are assigned to the categories C and P, but were such labels are randomised. The configuration which produces the

highest z-score $Z(\Gamma)$ is chosen as the best core-periphery partition of the system. In a slightly modified version, C_{ij} can be set to be equal to γ, with $0 < \gamma < 1$, if the pair of nodes i and j includes exactly one node of type C and one node of type P. However, despite in Fig. 6.1(a) block 2 is a block with only ones, as discussed again in Ref. [131] the specific value of the density of the links connecting a core node and a peripheral node is basically irrelevant to determine the goodness of a periphery structure, which is mainly based on maximising the difference in densities of block 1 and block 3. For such reason, the parameter γ is considered to be arbitrary. We also note that in this method the number of nodes N_c belonging to the core of the system is a second parameter to be determined a priori. The core-periphery structure of a small real-world citation network is shown in Fig. 6.1(b).

In large networks, testing all possible configurations in order to maximise $Z(\Gamma)$ and extract the core of a system, even with a predetermined size, is not feasible. Combinatorial optimization techniques such as simulating annealing can be used, but typically they make methods to find core-periphery structures in networks extremely slow, as they work taking into account the whole network structure.

Figure 6.2: Block structure of a network with community structure (a), core-periphery structure (b), global core-periphery structure with local community structure (c), and global community structure with local core-periphery structure (d). By permuting rows and columns of the adjacency matrix, one can see that (c) and (d) are equivalent [138].

Networks may in general display multiple cores, each one of them being formed by a separate group of central and densely connected units. We note that the two meso-scale structures considered in this thesis, namely community and core-periphery structures, are not mutually exclusive, but instead can coexist, as shown in Fig. 6.2. Conversely, bipartite networks are typically known to not have a core-periphery structure [137], as they have no clustering by default. At last, we remark that the models considered so far all produce a binary classification of nodes. Continuous models of core-periphery structure, where each node is assigned a value of coreness included between 0 and 1, have also been considered in the literature. A complete overview of different methods

and interesting alternatives can be found in Refs. [138, 139].

6.2 A local algorithm: the rich core

Methods to find cores in networks are usually quite complicated, parametric and computationally heavy. The Authors of Ref. [140] recently proposed a novel method which is fast, as it is based on local information only, non-parametric, and able to extract the so-called rich core of a network. The procedure to identify such rich cores consist of few simple steps. First, nodes are ranked in descending order of richness according to a chosen structural property ξ. Typically such property is chosen to be the degree k. Second, for each node i, its k_i links are divided into two groups: those towards richer neighbours which have a higher degree (connections of type '+', whose number we indicate as k^+), and those towards poorer neighbours with a lower degree (connections of type '-', whose number we indicate as k^-), such that

$$k_i = k_i^+ + k_i^-. \tag{6.2}$$

We then study the number of links towards richer nodes $k^+ = k^+(r)$ as a function of their rank in richness r. The node with the highest degree in a network has many connections, but they are all towards poorer nodes, i.e. $k^+(r = 1) = k^+(k = k_{\max}) = 0$. Conversely, the poorest node in the system will have only connections of type '+', but they will be very few, i.e. $k^+(r = N) = k^+(k = k_{\min}) = k_{\min}$. In general there exists a node with given rank r^*, such that k_r^+ reaches its maximum k_r^{+*}. By definition the rich core of the network is formed by all nodes with a rank which is better than or equal to r^*, whereas the remaining ones belong to the periphery. In the case of weighted networks, the degree k_i of a node i is replaced with its strength $s_i = \sum_j w_{ij}$, where $w_{ij} \geq 0$ describe the intensity of the connection between i and j, and k_i^+ becomess_i^+. The algorithm can also be easily generalised to the case of directed networks.

As it is based on local information only, profiling the rich core of a network is a much quicker procedure than other methods to identify core-periphery structures in networks, and it is particularly suited for large-scale systems. The algorithm is illustrated in Fig. 6.3 for the famous Zachary Karate Club network [141], and has been applied in a number of different contexts, including detecting the core of the network of collaborations and funding among universities in the United Kingdom [142]. In spite of its simplicity, this pragmatic way of defining a core was shown to be related to coupling the two concepts of escaping time of a random walker and the rich-club behavior of a network [140].

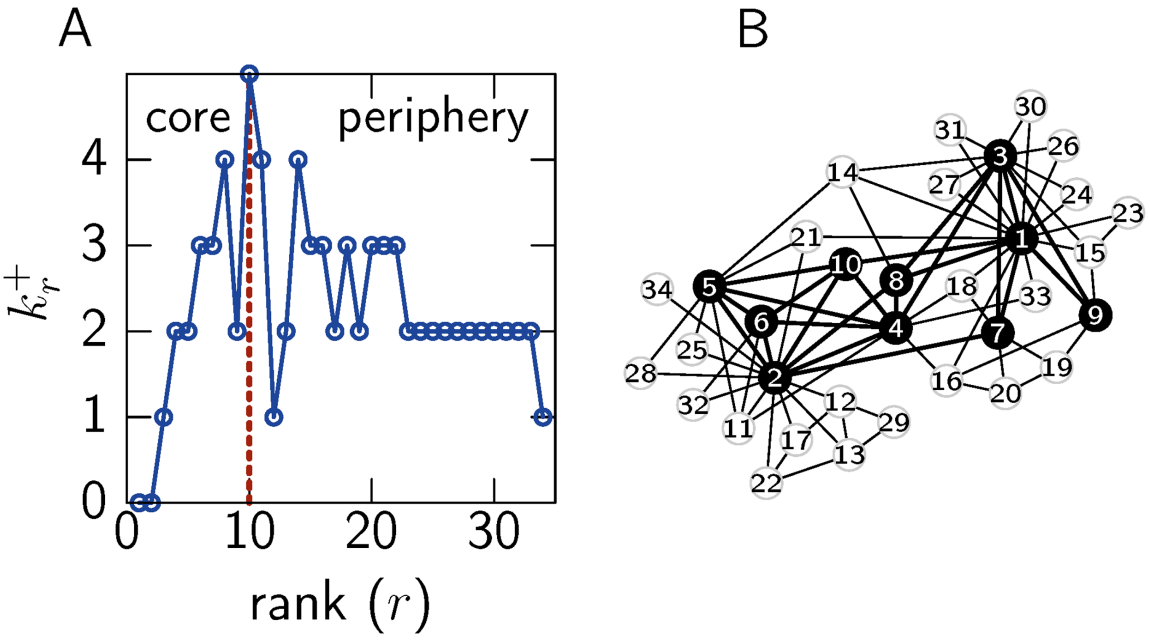


Figure 6.3: (a) Number of links k_r^+ towards richer nodes for each node r in the Zachary Karate club network [141]. The boundary of the core is marked by k_r^{+*} and indicated by the dotted line. (b) A graphical representation of the network with core (black) and periphery (white) nodes derived from the rich-core method [140].

We now evaluate the performance of such algorithm on a simple model inspired by the discussion of the previous Section [**?**]. Let us consider a stochastic block model where N_c nodes are labeled as core and $N_p = N - N_c$ as periphery. A connection between two core nodes (block 1) is generated with probability ρ_1, a connection between a core node and a periphery node (block 2) is generated with probability ρ_2, and a connection between two peripheral nodes (block 3) is generated with probability ρ_3. We choose the three probabilities such that $\rho_1 \geq \rho_2 \geq \rho_3$, in agreement with our previous discussion on the block structure of a network with a core-periphery structure. The stronger the block structure associated to the graph, i.e. the higher $\Delta\rho = \rho_1 - \rho_3$, the easier it is for the algorithm to recover the wanted core-periphery partition. Given the three probabilities, the expected total number of edges connecting two core nodes is $K_{cc} = \rho_1[(N_c - 1) * N_c/2]$, the expected total number of edges connecting two peripheral nodes is $K_{pc} = \rho_3[(N - N_c - 1) * (N - N_c)/2]$, and the expected total number of edges connecting a node from the core and a node from the periphery $K_{cp} = \rho_2[N_c * (N - N_c)]$. The total number of links is $K = K_{cc} + K_{cp} + K_{pp}$. It is also possible to compute the average degree of core nodes as

$$\langle k_c \rangle = \rho_1(N_c - 1) + \rho_2(N - N_c) \tag{6.3}$$

and that of the peripheral nodes as

$$\langle k_p \rangle = \rho_3(N - N_c - 1) + \rho_2(N_c). \tag{6.4}$$

The average degree of the system will then be equal to

$$\langle k \rangle = \frac{N_c \langle k_c \rangle + (N - N_c)\langle k_p \rangle}{N}. \tag{6.5}$$

In the following we fix the value of ρ_2 to a precise value, and study the system has a function of the relative density of the core and the periphery blocks. We start by considering a system with $N = 250$ nodes, of which $N_c = 50$ belonging to its core. We choose $\rho_2 = \bar{\rho} = 0.04$. For $\rho_1 = \rho_2 = \rho_3$, this condition fixes the average degree $\langle k \rangle \approx 10$. We are interested in studying the algorithm performance by varying the internal structure of the network, but keeping fixed the total number of connections. For such a reason, as we vary ρ_1, we can opportunely compensate by tuning the density of the periphery-periphery block and setting it equal to

$$\rho_3 = \frac{2}{(N - N_c) * (N - N_c - 1)} \Big(K - K_{cc} - K_{cp} \Big). \tag{6.6}$$

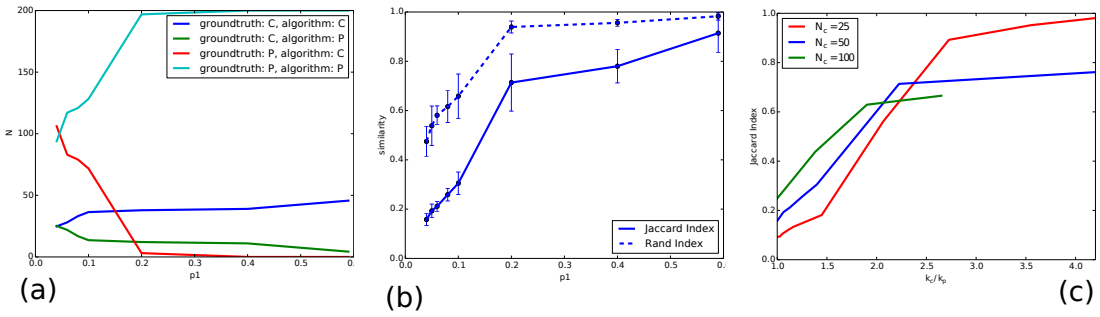


Figure 6.4: (a) Number of nodes assigned label C correctly (blue line) and wrongly (green line) identified by the algorithm, and number of nodes assigned label P correctly (light blue line) and wrongly (red line) identified by the algorithm, as a function of ρ_1. (b) Jaccard and Rand index between the pre-assigned groundtruth partition and the partition recovered by the rich core algorithm as a function of ρ_1. For the considered system size ($N = 250$ nodes, $N_c = 50$, $N_p = 200$) the algorithm appears to work extremely well for $\rho_1 > 0.2$. (c) Jaccard index between the groundtruth partition and the partition recovered by the algorithm for different core size as a function of the average degree of core and peripheral nodes k_c/k_p [**?**].

In Fig. 6.4(a) we show that, as ρ_1 grows, the total number of nodes originally labeled as cores and periphery correctly classified by the algorithm increases, whereas the number of mistakes diminishes [**?**]. Let us call n_{cc} the number of nodes labeled and identified by the algorithm as core, n_{pp} the number of nodes labeled and identified by the algorithm as periphery, n_{cp} the number of nodes labeled as core but identified as periphery (false negative) and n_{pc} the number of nodes labeled as periphery but identified as core (false positive). We evaluate the goodness of the extracted partition as a function of ρ_1 in terms of the Jaccard index

$$J = \frac{n_{cc}}{n_{cc} + n_{cp} + n_{pc}} \tag{6.7}$$

and the Rand index

$$R = \frac{n_{cc} + n_{pp}}{n_{cc} + n_{pp} + n_{cp} + n_{pc}}. \tag{6.8}$$

While R considers equally important a correct classification of core and peripheral nodes, J focuses on the first only. As shown in Fig. 6.4(b) both indexes increase as a function of ρ_1, even if R is higher than J. Whereas the increase is stronger for small values ρ_1, it becomes lighter for larger values, in particular after $\rho_1 = 0.2$. In Fig. 6.4(c) we show the Jaccard index between the partition found by the algorithm and the groundtruth for systems with cores of different sizes, such as $N_c = 25$, $N_c = 50$ and $N_c = 100$.

6.3 Rich cores in multiplex networks

In multi-layer networks nodes can in general play very different roles on different layers. For instance, a node which is extremely central in one layer might be very peripheral on a different one. Conversely, it is possible to have nodes that do not belong to the core of any layer individually, but because of their interplay are still quite important for the overall relational organisation of the multiplex network [**?**]. A first unsatisfactory solution to extract cores in multiplex networks consists in applying the single-layer method to the weighted overlapping network $\mathcal{O}$. However, in such case the specific role played by a node on each level is completely neglected. We remark that the meaning associated to the different types of links can be very different, as we saw in Chapter 3, where we considered the structural and the functional layers in the human brain.

Surprisingly, no method has been provided so far to profile core-periphery structures in multi-layer networks. Starting from the concept of rich core, we develop a method to detect multilayer core-periphery structure by accounting for interactions of different nature [**?**]. For each node i we compute its multiplex degree vector $\boldsymbol{k}_i = \{k_i^{[1]}, \ldots, k_i^{[M]}\}$. For each layer α, we then divide its links between those towards poorer and reacher nodes,

namely

$$k_i^{[\alpha]} = k_i^{-[\alpha]} + k_i^{+[\alpha]}. \tag{6.9}$$

Summing across the M layers, the total number of connections towards richer nodes, keeping track of the specificity of the node role at each level of the system, is

$$k^+_{i,\text{multiplex}} = \sum_{\alpha=1}^{M} k_i^{+[\alpha]}. \tag{6.10}$$

Differently from the corresponding variable $k^+_{i,\text{aggregated}}$ computed on the overlapping network $\mathcal{O}$ with the single-layer weighted algorithm, $k^+_{i,\text{multiplex}}$ does take into account the different contribution of the same nodes to the different layers of the system, rather than only its averaged value on the aggregated network. Nodes are then ranked in terms of overlapping degree $o_i = \sum_\alpha k_i^{[\alpha]}$, from highest to lowest. We look for the maximum value of $k^+_{i,\text{multiplex}}$, and all nodes from that with overlapping degree o_i higher or equal to that of the node corresponding to $k^+_{\text{multiplex,max,}}$ are part of the multiplex core, while the remaining ones are part of the multiplex periphery.

In order to illustrate how the algorithm works we evaluate such procedure on a novel multi-layer stochastic block models [**?**]. Other modelization of multiplex networks in terms of stochastic block models can be found in Refs. [143, 144]. As for the single-layer model, we consider a multiplex of $N = 250$ nodes, where in each layer α $N_c^{[\alpha]} = 50$ of them are labeled as core nodes. Once again we set the average degree such that $\langle k^{[\alpha]} \rangle = \langle k \rangle = 10$ and the densities of block 2 in the two layers are again set equal to $\rho_2^{[\alpha]} = \rho_2 = 0.04\ \forall \alpha$. The densities of the blocks 1 is set equal to $\rho_1^{[\alpha]} = \rho_1 = 0.2$, as this was evaluated as the minimum value on the single-layer case to achieve a sufficiently strong core-periphery block structure from the previous single-layer analysis. $\rho_3^{[\alpha]} = \rho_3$ is then computed through Eq. 6.6. We evaluate the algorithm over different multi-layer networks with $M = 2$ layers, where the core nodes of the two layers have different levels of node overlap. Core node overlap accounts for the fraction of core nodes in the first layer that are also core nodes in the second layer. In Fig. 6.5(a) we consider two layers whose core nodes completely overlap. Their block structure is equivalent (a, top), the degrees $k^{[1]}$ and $k^{[2]}$ are correlated and most of the nodes belonging to each core (a, medium) form the multiplex core of the system (a, bottom). In Fig. 6.5(b) we consider layers where half of the core nodes are shared with the other level of the system, and half are typical of each level. The block representation of the two layers is partially overlapping (b, top), and the nodes are spread uniformly over the plane according to the pairs $(k_i^{[1]}, k_i^{[2]})$ (b, medium). The multiplex core of the system (b, bottom) is formed by nodes which are part of the core on both layers and nodes which, despite being in the periphery in one of the two, score extremely high in the other one. At last, in Fig. 6.5(c)

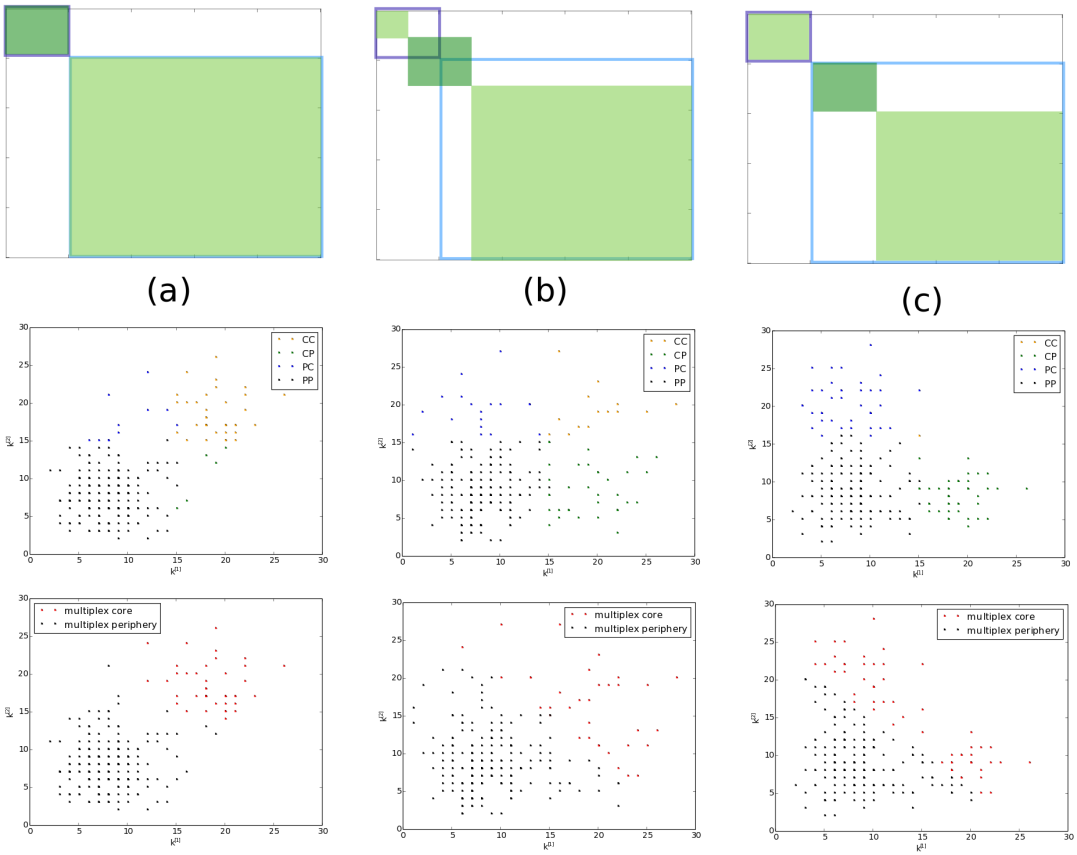


Figure 6.5: In panels (a),(b) and (c) we consider three multiplex networks with $M = 2$ and different levels of core node overlap. The block structures of the two layers of each multiplex are shown in the top row panels. The core block of layer 1 (its block 1) is delimited by a dark blue line, whereas the periphery block (its block 3) is delimited by a light blue line. The core block of layer 2 is coloured in dark green and the periphery block is coloured in light green. The core of layer 1 is fixed while the core of layer 2 is progressively shifted from (a) to (c) to decrease the node overlap from maximum to minimum. In the scatterplots we show each node as a function of the degree on the two layers according to a single-layer analysis (intermediate row: core in both layers, periphery in both layers, or mixed situation) and the multiplex algorithm (bottom row) [**?**].

we consider layers whose cores do not overlap (c, bottom), meaning that $k^{[1]}$ and $k^{[2]}$ are anti-correlated (c, medium). The multiplex core for such configuration is shown in (c, bottom).

6.4 Discussion

We note that the two considered functional forms to aggregate $\{k_i^{[1]}, \ldots, k_i^{[M]}\}$ and $\{k_i^{+[1]}, \ldots, k_i^{+[M]}\}$ represent the simplest possible choices. More in general, however, the multiplex richness μ_i (equal to o_i in the simplest case) will be a generic function f of the degree of a node at the different layers:

$$\mu_i = f(k_i^{[1]}, \ldots, k_i^{[M]}). \tag{6.11}$$

In an analogous way, the contribution towards richer nodes μ_i^+ (equal to $k_{i,\text{multiplex}}^+$ in the simplest case) will be given by

$$\mu_i^+ = g(k_i^{+[1]}, \ldots, k_i^{+[M]}), \tag{6.12}$$

where g is another generic function. In some contexts, exogenous information is available on the relative importance $c^{[\alpha]}$ of the M layers. Let us be both f and g two linear additive functions. The previous equations can be re-written to take into account the relative importance of the different layers so that

$$\mu_i = \sum_{\alpha=1}^{M} c^{[\alpha]} k_i^{[\alpha]} \tag{6.13}$$

and

$$\mu_i^+ = \sum_{\alpha=1}^{M} c^{[\alpha]} k_i^{+,[\alpha]}. \tag{6.14}$$

If no exogenous information on the layers is available, $c^{[\alpha]} = c = 1 \ \forall \alpha$ and μ_i and μ_i^+ correctly reduce to o_i and $k_{i,\text{multiplex}}^+$ respectively [**?**]. The presented method enables to extract a multiplex core from any type of multilayer datasets. We hope that in the future this technique will help unveil the core structures of many social, biological and man-made networks where the units interact across channels of different type.

Part II

Dynamics on multiplex networks

The first part of this thesis was devoted to set out the formalism to describe the structure of a multiplex network, which we take to be given either by eq. (1) or (2), and unveil the structure at the micro, meso and macroscale of real-world multiplex networks. In the second part of the thesis we investigate dynamical processes taking place on top of multiplex networks, from random walks to social and evolutionary dynamics, and how multiplexity often produces novel and unexpected *emergent multiplex behavior* which is not observed on any of the layers individually, nor on the corresponding aggregated network.

Let us start by proposing a general formalism to describe dynamics in multilayer systems [33]. We first define the dynamical state of the N nodes of the system. We assume here that the state of a node i at time t is described by a D-dimensional vector $\boldsymbol{\sigma}_i = \boldsymbol{\sigma}_i(t) \in \mathbb{R}^D$ whose components we take w.l.o.g. as real numbers. Notice that the dimensionality D of the state variable of each unit is not necessarily directly related to the number of layers M of the multiplex. In the simplest possible case the state of the node is identical across all layers, which means that $D = 1$ and $\boldsymbol{\sigma}_i$ is single-valued. Often, however, we want to model a case where a node's state can be different depending on the layer. In this more general case the global dimensionality of the node state as a combination of the dimensionality $d^{[\alpha]}$ at each layer, i.e. $D = \sum_{\alpha=1}^{M} d^{[\alpha]}$. When the state of the node on each layer has the same dimensionality d, the previous expression reduces to $D = M \times d$. In this way, the dynamical state of a multiplex network is determined once we give the $N \times D$-dimensional vector $\mathcal{S} = \{\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2 \ldots, \boldsymbol{\sigma}_N\}$.

In such most general form the equations governing the dynamics on a multiplex can be written as [33]:

$$\frac{d\mathcal{S}}{dt} = F(\mathcal{M}, \mathcal{S}) \tag{6.15}$$

for a continuous-time process, or

$$\mathcal{S}(t+1) = F(\mathcal{M}, \mathcal{S}(t)) \tag{6.16}$$

for a discrete-time process. We note that more complicated cases are also possible, such as coupling layers with both continuous and discrete time evolution, as in the case of Ref. [29].

In the equations given above the structure of the multiplex remains fixed, i.e. $\mathcal{M}(t) = \mathcal{M}(0)\ \forall t$, and only the states of the nodes $\mathcal{S}$ are changing over time. Now, it is also possible for the multiplex structure to evolve over time, $\frac{d\mathcal{M}}{dt} = G(\mathcal{M})$. When the evolution of the multiplex structure is influenced by a fixed state or topological property of the nodes, the previous equation can be generalised as $\frac{d\mathcal{M}}{dt} = G(\mathcal{M}, \mathcal{S})$. This naturally leads

to the most general case, where the dynamical evolution of the state of the nodes and the multiplex structure is interdependent. For this last case, we can write the dynamical equations as [33]

$$\begin{cases} \frac{d\mathcal{S}}{dt} = F(\mathcal{M}, \mathcal{S}) \\ \\ \frac{d\mathcal{M}}{dt} = G(\mathcal{M}, \mathcal{S}) \end{cases} \tag{6.17}$$

for continuous-time processes, and as

$$\begin{cases} \mathcal{S}(t+1) = F(\mathcal{M}(t), \mathcal{S}(t)) \\ \\ \mathcal{M}(t+1) = G(\mathcal{M}(t), \mathcal{S}(t)) \end{cases} \tag{6.18}$$

for discrete-time processes. We finally note that it is also possible to associate a dynamical state to the edges of the multiplex and consider the corresponding link dynamics. However, this topic is still in its infancy and will not be covered in this thesis.

This general formalism for dynamical processes on multiplex networks was introduced in Ref. [33], where Authors analyse selected examples from the multilayer literature where multiplexity produces genuinely novel emergent behavior. We note that there exists as well a number of thematic reviews on dynamics on multilayer networks, for instance on percolation [145], games [146], spreading processes [147] and based on the tensorial formalism [148]. In the next chapters of this thesis we focus on specific dynamics such as random walks, opinion dynamics, cultural dynamics and evolutionary game theory, closely presented as in the original papers, and investigate how the multilayer structure shapes and modifies the characteristic features of these processes.

Chapter 7

Biased random walks

In the realm of dynamical processes on networks [149] the simplicity and -still- the richness of random walks has attracted much attention in recent years [150, 151]. Random walks are the most simple way to explore a network using only local information, and the steady-state properties of a walk, including characteristic times, limiting occupation probability, and coverage, have tight relationships with the structure of the graph upon which the walk takes place [152, 153]. For this reason, random walks have also been successfully used as probes of network properties, with applications ranging from community detection [129, 154, 155] to taxonomy of real-world networks [156]. Moreover, specific flavours of random walks are widely used for the exploration of online social networks, information networks and the like.

A class of random walkers of particular interest is that of walkers whose motion is biased on the structural properties of the network [157]. In its simplest possible version, the considered *biased random walks* are Markov processes whose transition probability is a parametric function of the topological properties of the destination node. In this way, by tuning the parameters of the biasing function one can force the walk to preferentially visit, or avoid, nodes exhibiting high or low values of given topological descriptors, such as the degree, clustering or betweenness. In particular, degree-biased random walks have been used to define new centrality measures [158, 159], identify communities [160], and provide optimal exploration of a network using only local information [162]. It has also been found that the dynamics of degree-biased random walks is strongly affected by the presence of degree-degree correlations in the structure of the network [161, 163, 164], so that an appropriate choice of the structural bias can be used to perform efficient sampling of unknown networks.

In this chapter we study several ways in which random walks can be extended to

multi-layer networks, and we show how to devise appropriate ways to bias the walkers on the topological properties of the nodes at each layer in order to perform an efficient exploration of such systems [165]. We notice that random walks have also been applied to multi-layer networks to quantify the impact of failures in interconnected systems [166]. However, we will focus here on biased random walks and will investigate how the biasing function affects the dispersiveness of the walk and the steady-state occupation probability distribution [165]. The aim is to find walks which visit far away regions of a multiplex network within a relatively small number of steps, a property related to the dispersiveness of the walk, and, at the same time, guarantee that the probability for a walker to visit any node in the system is as close as possible to uniform, thus allowing to sample unknown graphs in an efficient way.

The presence of many interdependent layers allows to construct several classes of biased random walks, and in particular what we call *extensive walks* and *intensive walks*, where the difference between the two classes is in the dependence of the parameters of the biasing function on the number of layers of the system. In the former case, the biasing function depends on the structural properties of the destination node at all the layers of the system (thus, the number of parameters is extensive in the number of layers), while in the latter case the bias depends on intrinsically multiplex properties of the destination node, which do not depend explicitly on the number of layers of the network.

For both classes of biasing functions, we provide analytical closed forms for the long-time properties of the random walks, in terms of stationary probability distribution and entropy rate [167], and we study the effect of different structural properties, including the number of layers, the presence and sign of inter-layer degree correlations, the redundancy of edges across layers, the density of the multiplex and the heterogeneity of the degree distributions, on the steady-state behaviour of these walks. We find that all these properties have a remarkable effect on the maximal dispersiveness and on the steady-state occupation probability of biased random walks [165].

Finally, we study the diffusion properties of several real-world multiplex networks, namely the six continental airline transportation networks, and we show that in those cases the pressure to provide robust route alternatives has somehow hindered the overall diffusion properties of those systems [165].

7.1 General features of biased random walks

Let us consider a M-layer multiplex network of N nodes. In the following we assume the layers to be unweighted, but all the results can be easily extended to to the case of weighted multiplexes.

In general, a random walker on a multiplex is not constrained on a single layer and can exploit all the connections pointing out of the current node, at all layers. A synthetic – yet incomplete– description of the topology of a multiplex is provided by the overlapping adjacency matrix $\mathcal{O} = o_{ij}$, as introduced in Section 2.3, whose entries $o_{ij} = \sum_\alpha a_{ij}^{[\alpha]}$ account for the total number of connections between two nodes across all layers [15]. In particular, we consider the class of Markovian random walks defined by the transition probabilities:

$$\pi_{ji} = \frac{o_{ij} f_j}{\sum_j o_{ij} f_j}. \tag{7.1}$$

This set up is very general and allows for a variety of different motion rules. In fact, f_j can be either a function of some topological multiplex properties of the arrival node j, or an informative combination of some structural features of the destination node, measured at all or at a fraction of the layers. Notice that the unbiased random walk on the multiplex is obtained by setting $f_j = 1$, $\forall j \in V$. In this case a walker jumps out of node i by traversing one of the edges incident on i chosen with uniform probability and independently on the layer to which it belongs. It is worth noting that the use of the overlapping adjacency matrix $\{o_{ij}\}$ does not automatically make the walk in Eq. (7.1) equivalent to a random walk on the aggregated graph obtained by flattening all the layers in a single network. In general, if the biasing function f_j depends, either explicitly or implicitly, on the structural properties of node j in the multiplex network, the walk in Eq. (7.1) cannot be directly mapped on an equivalent walk on the aggregated graph.

Stationary probability distribution. Starting from the one-step transition probability given in Eq. 7.1 we derive closed forms for several asymptotic properties of the walk [165]. Following an approach similar to that used in Ref. [157], we now show that for any choice of the biasing function f_j the stationary probability distribution $\boldsymbol{p}^* = \{p_i^*\}$ of biased walks on multiplex networks can be analytically derived, under the hypotheses that *i)* the topological overlapping matrix $\mathcal{O}$ is primitive and that *ii)* f_j is a time-invariant function of any property of the destination node j. We start by considering the probability $p_{i\to j}(t)$ that a walker starting at node i will be found on node j after exactly t time steps:

$$p_{i\to j}(t) = \sum_{j_1, j_2, \ldots, j_{t-1}} \pi_{j_1,i} \times \pi_{j_2,j_1} \times \ldots \times \pi_{j,j_{t-1}}, \tag{7.2}$$

and the dual probability $p_{j\to i}(t)$:

$$p_{j\to i}(t) = \sum_{j_1,j_2,\ldots,j_{t-1}} \pi_{j_1,j} \times \pi_{j_2,j_1} \times \ldots \times \pi_{i,j_{t-1}}. \tag{7.3}$$

Comparing Eq. (7.2) with Eq. (7.3) and considering that the multiplex is undirected (i.e., $o_{ij} = o_{ji}$), we obtain

$$c_i f_i p_{i\to j}(t) = c_j f_j p_{j\to i}(t), \quad \forall i, j \in V \tag{7.4}$$

where $c_i = \sum_j o_{ij} f_j$. If the matrix $\mathcal{O}$ is primitive, then a stationary probability distribution exists and $\lim_{t\to\infty} p_{i\to j}(t) = p_j^*$, leading to the expression:

$$c_i f_i p_j^* = c_j f_j p_i^*. \tag{7.5}$$

By imposing the normalisation condition $\sum_j p_j^* = 1$ we finally get:

$$p_i^* = \frac{c_i f_i}{\sum_\ell c_\ell f_\ell}. \tag{7.6}$$

We notice that Eq. (7.6) is quite general, since it does not explicitly depend on the form of the biasing function or on the actual structure of each layer or of the topological overlapping matrix $\mathcal{O}$.

In many real-world application scenarios, e.g. in crawling the structure of online social networks, it is important to guarantee that for long enough times the walk will end up visiting all the nodes of the graph with the same probability. It is easy to prove that an unbiased random walk is not a good choice in this case, since its steady-state occupation probability distribution is proportional to the degree sequence, hence an appropriate bias should be used to avoid to visit hubs more frequently than poorly-connected nodes. In practice, it is not always possible to find a walk which produces exactly the same stationary occupation probability distribution for all the nodes, i.e. $p_i^* = \overline{p} = 1/N$, $\forall i$. However, one could instead require that the resulting stationary probability distribution, although not equal for all nodes, has the minimum possible variance. In particular, in the following we will focus on the normalised standard deviation of the stationary probability distribution:

$$\eta(p^*) = \frac{\sigma(p^*)}{\mu(p^*)} \tag{7.7}$$

where $\mu(p^*)$ and $\sigma(p^*)$ are the average and the standard deviation of $\boldsymbol{p}^*$, respectively. We will look for suitable combinations of the parameters of the walk that produce the smallest possible value of $\eta(p^*)$, corresponding to the maximum uniformity of the accessibility of the nodes attainable on a certain multiplex network.

Entropy rate. One classical measure to quantify the mixedness or dispersiveness of a walk on a graph is the entropy rate $h = \lim_{t\to\infty} S_t/t$ [167], where S_t is the Shannon entropy of the set of all the trajectories of length t generated from the walk rule, and h is the minimum amount of information necessary to describe the process [167]. In particular, $h = 0$ only if the walk generates exactly one possible trajectory, while h is maximum when all the trajectories are equiprobable. Intuitively, walks with a high mixedness can explore remote regions of a graph within a relatively small number of time-steps. This property is again desirable for the efficient exploration of unknown networks, where only local information is available. In particular, it is interesting to find a biasing function which guarantees that the walk does not remain trapped for too long in any region of the graph, and this is usually obtained by maximising the dispersiveness of the walk.

It is possible to show that the entropy rate for a Markov process can be expressed as

$$h = -\sum_{i,j} \pi_{ji} p_i^* \ln(\pi_{ji}), \tag{7.8}$$

which means that h depends only on the walk rule π_{ij} and on the stationary probability distribution [157]. By substituting the analytical expression for $\boldsymbol{p}^*$ given in Eq. (7.6) into Eq. (7.8) we get:

$$h = -\frac{1}{\sum_i c_i f_i} \left[\sum_i f_i \sum_j o_{ij} f_j \ln(o_{ij} f_j) - \sum_i f_i c_i \ln(c_i) \right]. \tag{7.9}$$

This expression has a natural upper bound, which reflects the case of random walks where all trajectories of the same length have equal probability. It is interesting to notice that, as shown by Burda et al. in Ref. [168], the maximal value of entropy rate attainable by any walk on a given single-layer graph depends on the structure of the graph, and in particular for an undirected graph it is equal to $\ln \lambda_{\max}$, where $\lambda_{\max}$ is the maximum eigenvalue of the adjacency matrix of the graph.

This result can be extended to the case of walks on multiplex networks as follows [165]. The total number of trajectories of length t generated by a walk defined as in Eq. (7.1) is equal to $N_t = \sum_{i,j} (\mathcal{O}^t)_{ij}$, where $\mathcal{O}^t$ is the t-th power of the overlapping adjacency matrix. In the limit of large t, we have

$$\widetilde{h}_{\max} = \lim_{t\to\infty} \frac{\ln N_t}{t} = \ln \lambda_{\max}, \tag{7.10}$$

where $\lambda_{\max}$ is now the maximum eigenvalue of the overlapping adjacency matrix $\mathcal{O}$ (this result is a direct consequence of the application of the power method). In general, the

maximal value of the entropy rate attainable with a particular motion rule will be smaller than or at most equal to $\widetilde{h}_{\max}$. Since obtaining high mixedness is a desirable property of a walk in many real-world applications, such as when searching for a given resource on a graph, in the following we will look for combinations of the parameters of different motion rules which can produce high values of h, to better approximate the corresponding value of $\widetilde{h}_{\max}$ allowed by the structure of the network.

Heterogeneous mean-field. In the particular case in which the bias function f_i depends only on the (vectorial) degree $\boldsymbol{k}_i = \{k_i^{[1]}, k_i^{[2]}, \ldots, k_i^{[M]}\}$ of node i, where by definition $k_i^{[\alpha]} = \sum_j a_{ij}^{[\alpha]}$ is the degree of node i at layer α, the expression for the stationary probability distribution can be considerably simplified [165]. Let us consider a heterogeneous mean-field, in which all the nodes belonging to the same degree class $\boldsymbol{k}$ are structurally indistinguishable. Under these assumption, and since f_i depends only on the degree, then for all the nodes i having the same degree $\boldsymbol{k}_i = \boldsymbol{k}$ we have $f_i = f_{\boldsymbol{k}_i} = f_{\boldsymbol{k}}$, but also $c_i = c_{\boldsymbol{k}_i} = c_{\boldsymbol{k}}$, and similarly:

$$p_{\boldsymbol{k}}^* = \frac{1}{C} f_{\boldsymbol{k}} c_{\boldsymbol{k}} = \frac{1}{C} f_{\boldsymbol{k}} \sum_{\boldsymbol{k'}} o_{\boldsymbol{k}\boldsymbol{k'}} f_{\boldsymbol{k'}} \tag{7.11}$$

where C is an appropriate normalisation constant to ensure that $\sum_{\boldsymbol{k}} p_{\boldsymbol{k}}^* = 1$. Eq. (7.11) means that all the nodes in the same degree class will have the same steady-state probability of being visited by the walk. Notice that $o_{\boldsymbol{k}\boldsymbol{k'}}$ is the expected number of edges connecting two nodes whose multiplex degree is respectively equal to $\boldsymbol{k}$ and to $\boldsymbol{k'}$. If we assume that there are no edge correlations, i.e. that the probability of having $a_{ij}^{[\alpha]} = 1$ does not depend on the probability of having $a_{ij}^{[\beta]} = 1$ for all the possible $\beta \neq \alpha$, then we can write:

$$p_{\boldsymbol{k}}^* = \frac{1}{C} f_{\boldsymbol{k}} \sum_{\boldsymbol{k'}} f_{\boldsymbol{k'}} \sum_{\alpha=1}^{M} k^{[\alpha]} P(k'^{[\alpha]} | k^{[\alpha]}) \tag{7.12}$$

since the expected number $o_{\boldsymbol{k}\boldsymbol{k'}}$ of edges between a node with degree $\boldsymbol{k}$ and a node with degree $\boldsymbol{k'}$ is actually equal to the sum of the expected number of edges connecting these two nodes at each of the M layers (we indicate by $k'^{[\alpha]}$ the degree at layer α of a node whose vectorial degree is equal to $\boldsymbol{k'}$). If we additionally assume that there are no intra-layer correlations, then:

$$P(k'^{[\alpha]} | k^{[\alpha]}) = q_{k'^{[\alpha]}} = \frac{k'^{[\alpha]} P(k'^{[\alpha]})}{\langle k'^{[\alpha]} \rangle} \tag{7.13}$$

where $P(k'^{[\alpha]})$ is the degree distribution at layer α. In the end we find:

$$p^*_{\boldsymbol{k}} = \frac{1}{C} f_{\boldsymbol{k}} \sum_{\boldsymbol{k'}} f_{\boldsymbol{k'}} \sum_{\alpha=1}^{M} \frac{k^{[\alpha]} k'^{[\alpha]} P(k'^{[\alpha]})}{\langle k'^{[\alpha]} \rangle}. \tag{7.14}$$

This expression for $p^*_{\boldsymbol{k}}$ is quite general, and in particular it is valid even in the presence of inter-layer degree-correlations [22]. Since the heterogeneous mean-field discards intra-layer and edge correlations, which usually contribute to hinder the dispersiveness of a walk, Eq. (7.14) can be readily plugged into the expression of the entropy rate in Eq. (7.8) to obtain an estimate of the maximum value of h attainable with a given biasing function on a multiplex network with an assigned multiplex degree sequence $\{\boldsymbol{k}_i\}$.

7.2 Classes of biased random walks

The introduction of a biasing function in the motion rule is mainly motivated by the necessity to obtain an exploration of the graph which is more efficient, i.e., faster with respect to the time needed to visit all the nodes, or more homogeneous, i.e., avoiding heterogeneities in the stationary distribution probability, in order to explore with the same probability each node of the graph. In single layer networks these two aims are in general antithetical. For instance, a biasing function which maximises the mixing of the walk (corresponding to higher values of entropy rate) usually produces a quite heterogeneous stationary occupation probability, mainly due to the fact that a better mixing is obtained by exploiting the central role played by hubs. High values of h are usually achieved in a single-layer uncorrelated graph by a degree-biased walk $\pi_{ji} \sim k_j^b$ with $b = 1$, and in general with a bias $b > 0$ in graphs with non-trivial degree-degree correlations [157]. On the other hand, a uniform stationary occupation probability is obtained by using $\pi_{ji} \sim k_j^b$ with $b = -1$ in uncorrelated graphs, and in general by a value of $b < 0$ for graphs with degree-degree correlations, which corresponds to forcing the walkers to preferentially move towards poorly connected nodes [164].

The richness of multi-layer networks allows the exploration of more complex biasing functions and, as we will show in the following, usually produces quite interesting dynamics. The reason of such richness is that the multiplex degree of a node i is a vectorial rather than a scalar quantity, a fact that allows to construct several degree-based biasing functions. In the following we present two particular classes of such biasing functions, which we call *extensive* and *intensive* biases, respectively [165].

Extensive bias functions. We call *extensive* those walks whose motion rule depends on a function of the degrees of the destination node at each of the M layers. A first

example is that of *additive* degree-biased walks, defined by transition probabilities of the form:

$$\pi_{ji} \propto \sum_{\alpha=1}^{M} (k_j^{[\alpha]})^{b_\alpha} \tag{7.15}$$

where $b_\alpha \in \mathbb{R}$ is the bias exponent associated to layer α. Another example is that of *multiplicative* degree-biased walks, whose transition probability takes the form:

$$\pi_{ji} \propto \prod_{\alpha=1}^{M} (k_j^{[\alpha]})^{b_\alpha}. \tag{7.16}$$

We named these walks "extensive" since the number of free parameters in the motion rule, namely the exponents b_α, increases with the number of layers M. This peculiar property of extensive walks allows for a fine-grained setting of the bias in order to avoid nodes whose replicas on each of the M layers belong to a specific degree class. For instance, in the case of a two-layer multiplex, if we set $b_1 > 0$ and $b_2 < 0$ then the walkers will preferentially move towards node having, at the same time, high degree on layer 1 and low degree on layer 2. It might sometimes be desirable for a walker to have such sophisticated motion rules. An example is that of multiplex collaboration networks, in which nodes are scientists and layers represent co-authorship patterns in different fields. In that case, we might use an appropriately biased multiplex random walk which prefers to move towards nodes having a higher degree in a particular field, whose stationary probability distribution will represent a measure of the relative importance of each author in that field.

However, having a number of parameters which scales with the number of layers is not always a desirable property, especially if one wants to tune these parameters in order to obtain a walk with certain dynamical properties (e.g., either in terms of stationary probability or in terms of entropy rate). This problem is efficiently solved by intensive bias functions.

Intensive bias functions. We call *intensive* those multiplex walks whose motion bias depends on one or more intrinsically multiplex properties of the destination node. In the following we will focus on the intensive walk whose transition probability reads:

$$\pi_{ji} = \frac{o_{ij}(o_j^{b_o}\mathcal{P}_j^{b_p})}{\sum_\ell o_{i\ell}(o_\ell^{b_o}\mathcal{P}_\ell^{b_p})} \tag{7.17}$$

where $o_j = \sum_\alpha k_j^{[\alpha]}$ is the overlapping degree of node j and $\mathcal{P}_j$ is the multiplex participation coefficient of j, as defined in Section [15]. We notice that by considering $o_\bullet$ and $\mathcal{P}_\bullet$

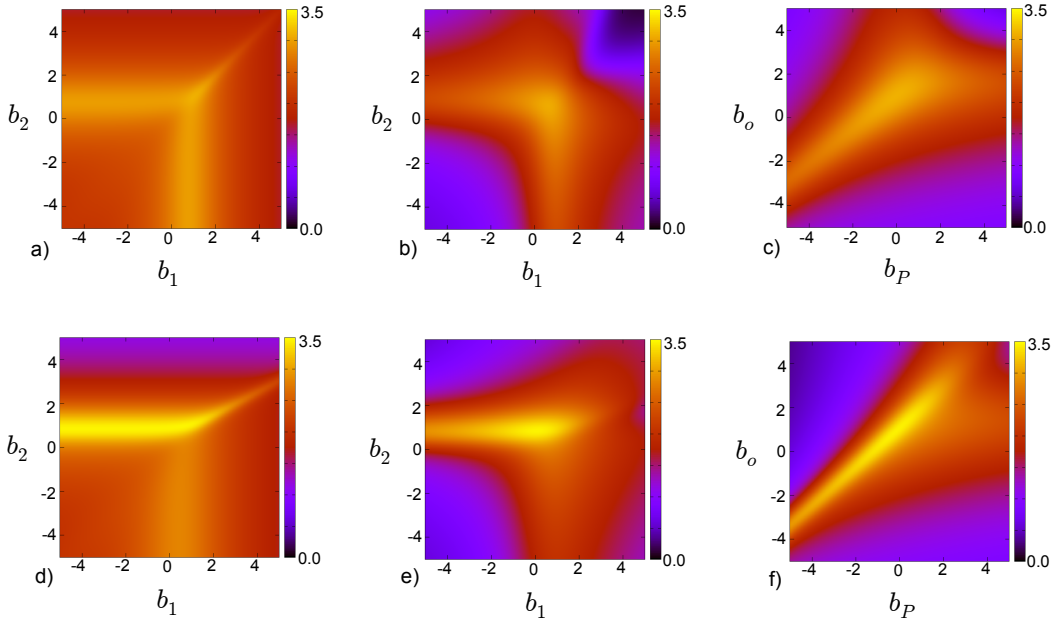


Figure 7.1: Heat-maps of the value of entropy rate h of different multiplex biased walks as a function of the parameters of the biasing function. The panels correspond, respectively, to additive [right, (a) and (d)], multiplicative [middle, (b) and (e)] and intensive walks [left, (c) and (f)] on uncorrelated duplex networks (in the top panels the two layers have the same power-law degree distribution $P(k) \sim k^{-\gamma}$ with $\gamma = 2.5$, while in the bottom panels the two layers have power-law degree distributions with different exponents, namely $\gamma_1 = 2.2$ and $\gamma_2 = 2.7$. In general, the maximum of h is obtained for positive values of the two bias parameters, corresponding to extensive walks which move preferentially towards nodes having high degrees on both layers, and to intensive walks whose motion rule is biased towards truly multiplex nodes [165].

we are effectively using information about the *distribution* of the edges of the destination node across the layers. In particular, for fixed number of layers M, o_i is proportional to the average of the distribution defined by $\boldsymbol{k}_i = \{k_i^{[1]}, k_i^{[2]}, \ldots, k_i^{[M]}\}$, while $\mathcal{P}_i$ gives information about the homogeneity of $\boldsymbol{k}_i$, with $\mathcal{P}_i \sim 1$ if $k_i^{[\alpha]} \simeq \frac{1}{M}\sum_\beta k_i^{[\beta]}\ \forall \alpha$ (i.e., if node i has roughly the same degree at all layers) and $\mathcal{P}_i \sim 0$ if almost all the edges of node i lie on just one layer.

We notice that when $b_o > 0$ the walkers will preferentially move towards hubs, while for $b_o < 0$ they tend to visit the poorly connected nodes more often. Similarly, for positive values of b_p the walkers will preferentially move towards *truly multiplex nodes*, i.e. nodes whose distribution of edges across the M layers is more homogeneous, while for

$b_p < 0$ the walkers prefer to move towards *focused nodes*, i.e. those having the majority of their connections in just one or a few of the M layers [15]. In general, by tuning the two parameters b_o and b_p we can obtain a rich variety of different walks. For instance, for $b_o > 0$ and $b_p > 0$, the walkers will be attracted by truly multiplex hubs (i.e., nodes with many links, almost equally distributed across the layers). Conversely, when $b_o > 0$ and $b_p < 0$ focused hubs are visited often and multiplex poorly connected nodes are strongly avoided, and so forth. The unbiased multiplex walk is recovered for $b_p = b_o = 0$.

The most interesting characteristic of the intensive walk defined by Eq. (7.17) is that the number of free parameters is fixed and does not scale with the number of layers, as instead happens for extensive walks. We will show in the following that intensive walks usually perform at least as well as extensive walks, e.g. with respect to the maximisation of entropy rate or to the minimisation of heterogeneity in the stationary occupation probability distribution.

It is worth noting that in the case of a duplex, i.e. when $M = 2$, even if the number of biasing parameters in intensive and extensive walks is the same, their effect on the motion of the walkers is different. Differently from b_1 and b_2, intensive biases do not allow to bias the walkers towards nodes with given properties in a particular layer but always consider intrinsically multiplex features, such as their total number of connections and their heterogeneity.

In order to explore the differences in the dynamical properties (i.e., the entropy rate h and the normalised standard deviation of the stationary occupation probability distribution $\eta(p^*)$) of biased multiplex walks, in the top panel of Fig. 7.1 we report the values of h obtained by additive, multiplicative and intensive random walks as a function of the two bias exponents in a two-layer multiplex network whose layers have the same average degree $\langle k \rangle$ and power-law degree distributions $P(k) \sim k^{-\gamma}$ with $\gamma = 2.5$, with no inter-layer correlations and no edge overlap. We note that the results obtained for different values of the exponent γ of the power-law degree distribution are comparable to those shown in Fig. 7.1 and Fig. 7.2, and are not reported for brevity. We notice that also in this simple case the three walks have remarkably different behaviours. In particular, the additive walk exhibits a relative small sensitivity to the values of the biasing exponents, which results in smaller variations of h. In fact, there is a large region of b_1 (i.e. $0 < b_1 < 2$) within which the entropy rate is almost constant and not very different from the absolute maximum for a relatively large range of values of the other exponent b_2, i.e. $-5 < b_2 < 2$ (the same reasoning is valid for $0 < b_2 < 2$ and $-5 < b_1 < 2$, due to the symmetry of the additive bias function).

Conversely, the picture is much richer and less trivial in the case of multiplicative

and intensive walks, for which the maximum of h is obtained for a relatively small range of parameters, usually corresponding to positive exponents. We obtain slightly different results when we consider two layers with different power-law degree distributions $P(k^{[1]}) \sim \left(k^{[1]}\right)^{\gamma_1}$ and $P(k^{[2]}) \sim \left(k^{[2]}\right)^{\gamma_2}$, namely with exponents $\gamma_1 = 2.2$ and $\gamma_2 = 2.7$ respectively. In this case, the symmetry in the additive and multiplicative phase diagrams is broken, and the maximum values of h are found by biasing the walk towards nodes with high degree on both layers, with a higher biasing exponent on the degree of the second layer, which has a more homogeneous degree distribution. Also the phase diagram for the intensive walk is modified, with the line of maximum values becoming thinner.

Similar considerations hold for the phase diagram of $\eta(p^*)$, reported in Fig. 7.2. In this case, the minimum variance (yielding a more homogeneous exploration of nodes) is obtained for negative values of the two bias exponents. Moreover, the phase diagram exhibits quite small variations in the case of additive walk, while we observe more heterogeneity in the case of multiplicative and intensive walks. Again, the symmetry of the phase diagrams of the extensive walks is broken when pairs of layers with different power-law exponents γ_1, γ_2 are considered, with the region $b_2 > b_1$ showing greater variations than for $b_2 < b_1$. Qualitatively similar differences can be obtained with asymmetric layers with respect to other statistical properties, such as density.

All the results for synthetic networks, both in the current and following sections, have been obtained for layers with $N = 10^4$ nodes and averaged over 1000 realisations.

7.3 How the structure of a multiplex affects the walk

In this section we illustrate how the structure of the multiplex network affects the maximal entropy rate and the minimum heterogeneity of the stationary occupation probability distribution achievable in the system.

We focus on five structural aspects, namely *i)* the presence and sign of inter-layer degree-degree correlations, *ii)* the existence of edge overlap across layers, *iii)* the number M of layers of the multiplex, *iv)* the power-law exponent γ of the degree distribution of the layers, and *v)* their density, measured through the average degree $\langle k \rangle$. Since our focus is on the construction of efficient walks (in terms of maximal dispersiveness and of homogeneity of the stationary occupation probability) the parameters of interest in all the cases are the overall maximum value of entropy rate, denoted by $h_{\max}$, and the minimum value of the normalised standard deviation, denoted by $\eta_{\min}$, obtained by extensive and intensive biased random walks as a function of the biasing parameters.

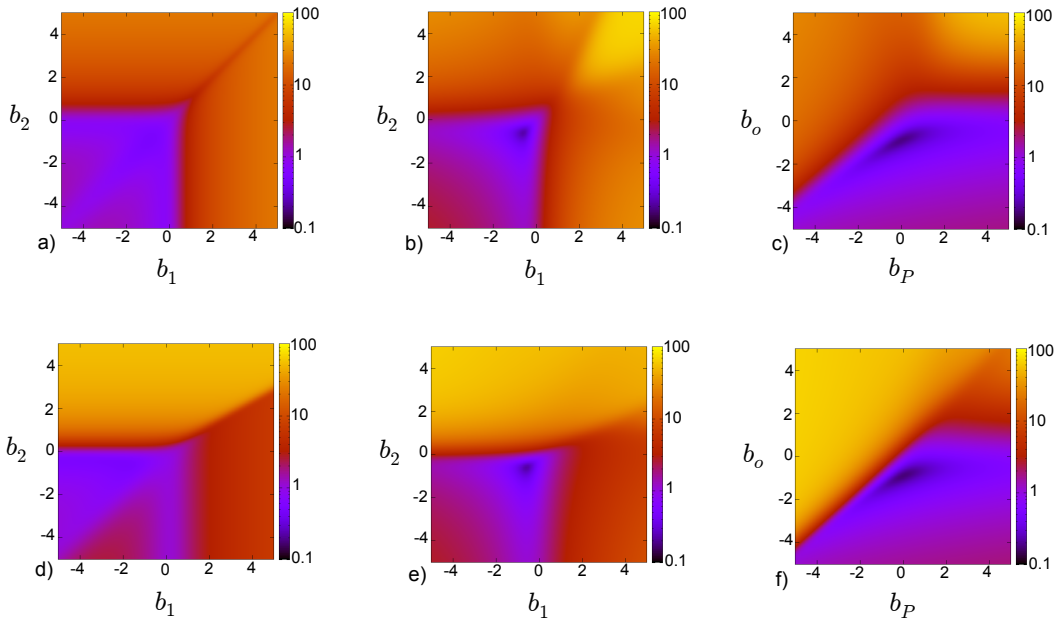


Figure 7.2: Heat-maps of the normalised standard deviation of the stationary occupation probability distribution $\eta(p^*)$ of different multiplex biased random walks. Legend as in Fig. 7.1. In extensive walks, the minimum of η is always attained for negative values of the two exponents, while in intensive walks the minimum of η is obtained for $b_o < 0$ and $b_P \simeq 0$, meaning that walkers tend to preferentially move towards nodes with small degrees on both layers [165].

Effect of inter-layer degree correlations. In this paragraph we recall the Spearman rank correlation coefficient ρ between layer α and layer β introduced in Section 2.2. We remind that the coefficient ρ takes values in $[-1, 1]$, so that $\rho = 1$ if the two degree sequences are perfectly correlated (meaning that a hub at layer α is also a hub at layer β), while $\rho = -1$ when the two degree sequence are perfectly anti-correlated, i.e. when a hub on layer α is always a poorly connected node on the other layer, and viceversa. Intermediate positive (negative) values of ρ indicate weaker positive (negative) inter-layer correlations, while $\rho \simeq 0$ when the two degree sequences are uncorrelated.

In Fig. 7.3(a) we report the plot of $h_{\max}$ and $\eta_{\min}$ for extensive and intensive walks on two-layer multiplex networks with same average degree and power-law degree distributions $P(k) \sim k^{-\gamma}$ with $\gamma = 2.5$, for different levels of inter-layer degree correlations [165]. As made evident by the figure, intensive walks usually perform at least as well as extensive walks with respect to both maximisation of entropy and minimisation of the heterogeneity of the stationary occupation probability distribution. This suggests that, aside from the actual differences in the phase space, intensive walks are able to span the same range

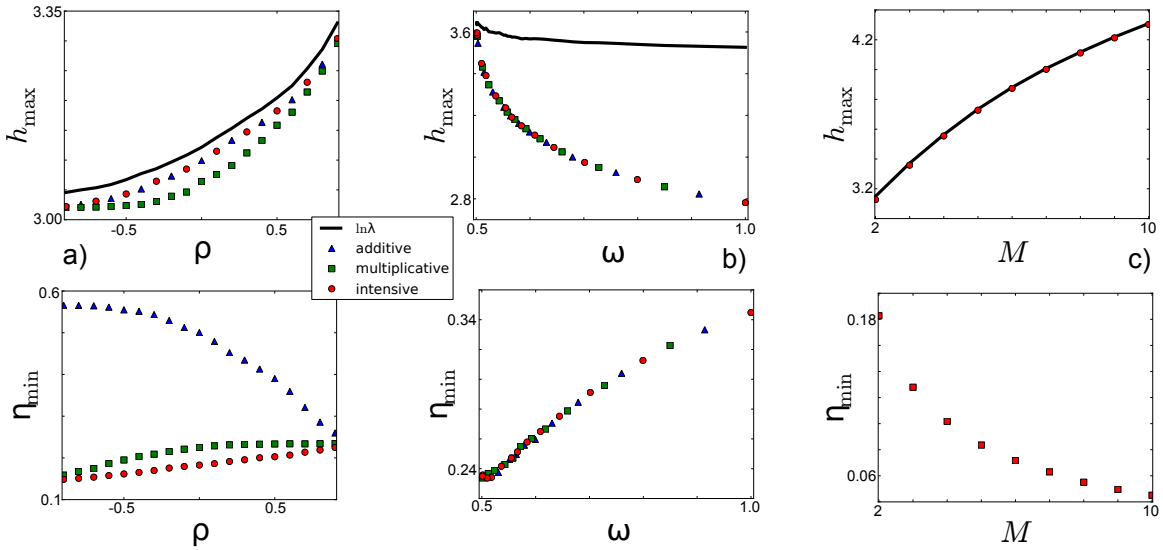


Figure 7.3: Values of $h_{\max}$ (top panels) and $\eta_{\min}$ (bottom panels) as a function of the the inter-layer degree correlation coefficient ρ (a), the average edge overlap ω (b) and the number of layers M (c), respectively for additive (triangles), multiplicative (squares) and intensive (circles) walks. For the entropy rate, we also show the value of $\widetilde{h}_{\max} = \ln \lambda$ corresponding to the maximum entropy random walk (solid line). (a) For all walks, $h_{\max}$ is an increasing function of the inter-layer degree correlation coefficient ρ, and provides a very good approximation of the maximum theoretical entropy rate $\widetilde{h}_{\max}$. Notice that intensive walks perform at least as well as the extensive ones. (b) As the overlap increases, the estimates of $h_{\max}$ obtained by the biased walks become less precise, while $\eta_{\min}$ increases as a function of ω. (c) $h_{\max}$ increases and $\eta_{\min}$ decreases as a function of M. In this case we only performed simulations for intensive walks [165].

of values of entropy and $\eta(p^*)$ by using only two parameters, irrespective of the actual numbers of layers in the multiplex.

Effect of edge overlap. We now investigate the impact of the presence of edge overlap on the long-term dynamics of extensive and intensive walks [165]. Given the multiple possible definition for the overlap of each edge (see Section 2.3), we consider here $\omega_{ij} = \frac{1}{M}\sum_{\alpha=1}^{M} a_{ij}^{[\alpha]}$. The edge overlap of a multi-layer network is then computed as the average of ω_{ij} over all the node pairs for which $o_{ij} \neq 0$ (i.e., for all pairs of nodes which are connected by at least one edge):

$$\omega = \frac{1}{\sum_{i,j}(1-\delta_{o_{ij},0})}\sum_{ij}\omega_{ij} = \frac{1}{2K}\sum_{ij}\omega_{ij} \tag{7.18}$$

where K is the number of pairs of nodes which are connected in at least one of the M layers. Notice that the average edge overlap ω is equal to 1 only if all the M layers are identical, while $\omega = 1/M$ when every edge is present on exactly one of the M layers.

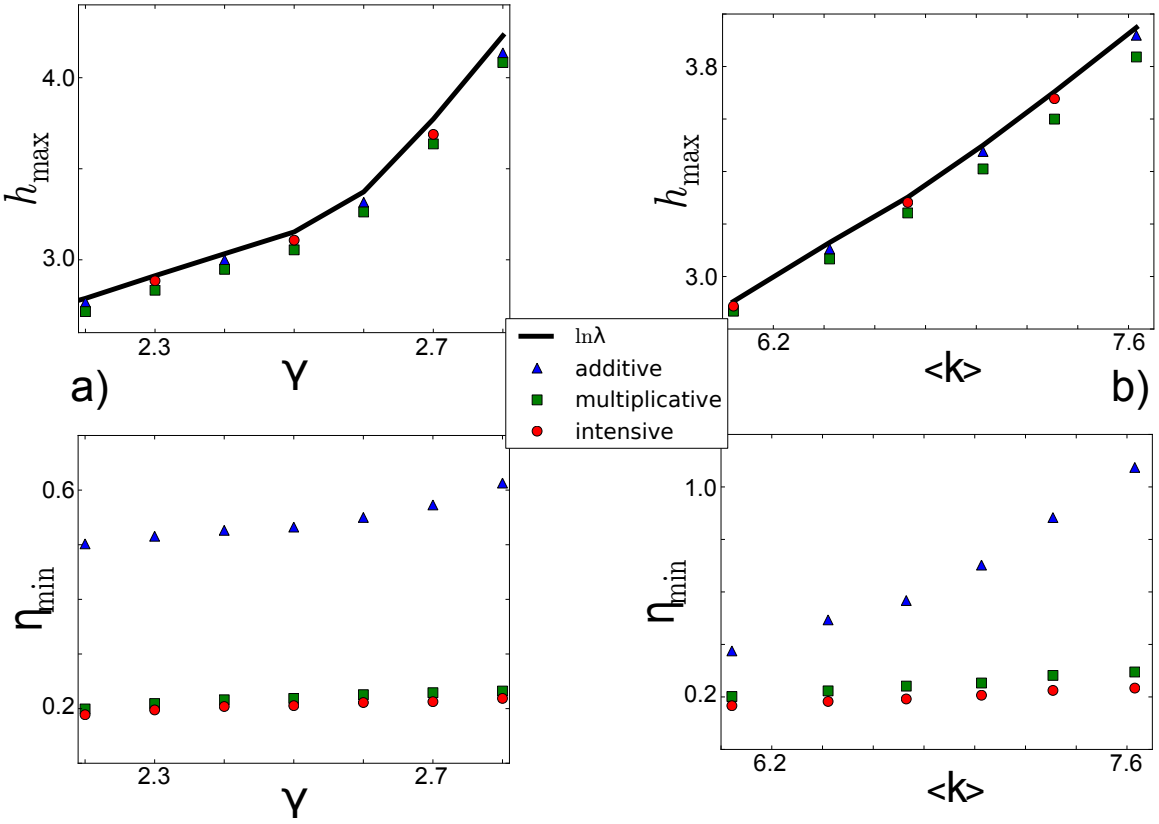


Figure 7.4: Values of $h_{\max}$ (top panels) and $\eta_{\min}$ (bottom panels) as a function of the exponent γ of the the power-law distribution of each layer (a) and of the average degree $\langle k \rangle$ (b), respectively for additive (triangles), multiplicative (squares) and intensive (circles) walks. For the entropy rate, we also show the value of $\widetilde{h}_{\max} = \ln \lambda$ corresponding to the maximum entropy random walk (solid line). As shown, for all walks $h_{\max}$ appears to increase as a function of both γ and $\langle k \rangle$. Smaller variations are also found in the values of $\eta_{\min}$ [165].

We started from two-layer multiplex networks obtained by coupling identical layers (thus having edge overlap equal to 1) with power-law degree distributions $P(k) \sim k^{-\gamma}$ with $\gamma = 2.5$, and then we obtained multiplex networks with prescribed values of edge overlap by rewiring a certain percentage of the edges of one of the two layers in order to maintain the degree sequence unaltered. Notice that by construction the resulting multiplex networks have maximally positive inter-layer degree correlations (i.e., $\rho = 1$). As shown in Fig. 7.3(b), $\eta_{\min}$ becomes higher as ω increases, meaning that higher values

of edge overlap correspond to a more heterogeneous stationary state probability distribution. Conversely, $h_{\max}$ decreases with ω, in accordance with the fact that higher edge overlap tends to hinder the dispersiveness of the walk, since a smaller number of distinct walks can originate from each node. Summing up, multiplex networks having smaller values of edge overlap are overall preferable in order to maximise the dispersiveness of the walk and to obtain a more homogeneous stationary occupation probability. In other words, a small edge overlap guarantees a more effective exploration of a multiplex network and, at the same time, a more homogeneous distribution of the probability of visiting each node.

Effect of the number of layers. It is also interesting to study how the dynamical properties of intensive walks change when the number of layers M is progressively increased [165]. To this aim, we constructed multiplex networks with different number of layers, with no inter-layer degree correlations and negligible edge overlap, where all the layers had power-law degree distributions $P(k) \sim k^{-\gamma}$ with $\gamma = 2.5$. As shown in Fig. **??**(c), $h_{\max}$ is an increasing function of M, while $\eta_{\min}$ decreases as the number of layers grows. In general, the addition of layers in absence of inter-layer correlation flattens the structural differences among the nodes of the multiplex, and provides better dispersiveness and less heterogeneity in the occupation probability distribution.

Effect of the heterogeneity of the degree distribution. We investigate here how the heterogeneity of the the degree distribution of each layer affects $h_{\max}$ and $\eta_{\min}$ [165]. To this aim, we considered pairs of uncorrelated layers with the same power-law degree distribution $P(k) \sim k^{-\gamma}$ for different values of γ, maintaining fixed the average degree of the networks $\langle k \rangle$. The plots in Fig. 7.4(a) confirm that both $h_{\max}$ and $\eta_{\min}$ grow as γ increases, i.e. as the degree distribution of the layers becomes more homogeneous. We notice though that the variation in $\eta_{\min}$ appears to be relatively smaller, especially for multiplicative and intensive walks. This result can be explained by considering that dispersiveness is favoured by more homogeneous degree distributions. Layers with different power-law exponents γ_1 and γ_2 have been considered in the previous section.

Effect of layer density. Finally, we focus on the effect of layer density, measured through the average degree of the layers $\langle k \rangle$ [165]. Once again we report here the case of uncorrelated layers with power-law exponent $\gamma = 2.5$, but similar results have been obtained for other values of γ. As shown in Fig. 7.4(b), both $h_{\max}$ and $\eta_{\min}$ increase as a function of $\langle k \rangle$. Layers with different average degrees $\langle k^{[1]} \rangle$ and $\langle k^{[2]} \rangle$ break the symmetry of the phase diagrams for h and η qualitatively in a similar way as pairing layers with different power-law exponents.

Summing up, the analysis of the impact of structural properties on the values of

$h_{\max}$ and $\eta_{\min}$ attainable on a multiplex network confirms that positive inter-layer degree correlations, small edge overlap, large number of layers, and more homogeneous layers all concur towards allowing biased random walks with nearly-optimal dispersiveness and closely-to-homogeneous steady-state visiting probability. In other words, a multiplex network with a large number of layers and small edge overlap, where nodes have roughly the same number of links at all layers, can be explored ways more efficiently than a similar multiplex network where nodes have disassortative inter-layer correlations and edges are redundant across layers.

In the following section we show that the multiplex airline transportation networks of all the six continents have evolved towards a structure which provides a good trade-off between efficient exploration and robustness.

7.4 Applications to real-world airline transportation networks

As an application, we study here the dynamical properties of multiplex biased walks on a set of real-world systems [165]. Namely, we apply our framework to the six continental airline transportation networks introduced in Chapter 1. As shown in Table 7-A, all such multiplex networks consist of a relatively high number of layers. For this reason, we will use intensive walks to compute the maximal entropy rate $h_{\max}$ and the minimum value of the standard deviation of the stationary distribution $\eta_{\min}$. In Table 7-A, we also report for each multiplex the average number of layers $M \times \omega$ where each edge exists, the theoretical upper value of entropy rate $\ln \lambda$, and the values of $h_{\max}$ and $\eta_{\min}$ obtained by optimising intensive walks.

Multiplex	N	M	$M \times \omega$	$\ln \lambda$	$h_{\max}$	$\eta_{\min}$
Africa	238	84	1.57	3.36	2.20	1.36
Asia	795	213	2.16	4.96	3.52	1.17
Europe	594	174	1.55	4.60	3.76	1.06
North America	1029	143	1.56	4.70	3.75	1.35
Oceania	261	37	1.52	3.71	2.39	2.00
South America	300	58	1.81	3.66	2.59	1.08

Table 7-A: Structural properties of the six continental airline transportation systems. For each multiplex, we report the number of nodes N, the number of layers M, the average number of layers in which an edge exists $M \times \omega$, the theoretical upper value of entropy rate $\ln \lambda_{\max}$ and the extremal values $h_{\max}$ and $\eta_{\min}$ obtained by optimising intensive walks [165].

We notice that the efficiency of a transportation system is usually measured in terms of the accessibility of the locations it serves. In particular, in an ideal transportation system it should be easy to travel between any pair of far-apart regions of the network, mostly irrespective of where exactly those locations are located. Now, discarding the cost associated to the distance between the nodes of an airline transportation network, high accessibility can be obtained by guaranteeing that a traveller can reach remote locations in the system without large effort, in terms of number of interchanges, and that all locations can be visited with comparable effort. We have seen that in the language of random walks these two criteria correspond, respectively, to the maximisation of dispersiveness and to the minimisation of the standard deviation of the visiting probability.

Hence, we can ask whether the six continental air transportation systems can guarantee a good level of navigability, i.e. an optimal trade-off between dispersiveness and homogeneity of the visiting probability. We reckon that a more informative analysis of the efficiency of these systems would require more detailed information about the actual patterns of trips travelled by passengers, the cost associated to each route, the presence of non-Markovian effects (people often come back to their original place at the end of a trip), the non-stationarity of the system due to seasonality, etc.. However, we argue that biased random walks can still provide useful, yet coarse-grained, information about the overall navigability of those systems. Since we cannot modify the degree distributions of each of the layers, or the patterns of inter-layer correlations, or the actual number of layers in each continental air transportation system, we focus here in particular on the effects of edge overlap.

In the previous section we showed that networks with high edge overlap ω achieve lower maximal values of dispersiveness of the walk and larger heterogeneity of the equilibrium occupation probability distribution. When two nodes are connected by more than one edge, indeed, from a dynamical point of the view some connections are wasted, since redundant links do not allow for new paths in the network. However, their redundancy might often be important for a transportation system, since it makes specific connections more robust to single link failures. It is not unrealistic to assume that multi-layer transportation systems from the real-world have developed by satisfying a trade-off between the necessity to provide, at the same time, high diffusivity together with reasonable levels of robustness, a common problem in spatial networks [166, 169].

Because of the large heterogeneity in the size and number of layers of the six continental transportation systems, it is necessary to introduce some kind of normalisation which allow to compare the results observed in different systems. In order to test the effect of edge overlap on the diffusion properties of real-world systems, for each of the

six multiplex networks we computed the z-score of the average edge overlap:

$$z(\omega) = \frac{\omega - \langle\omega\rangle}{\sigma(\omega)}, \tag{7.19}$$

where $\langle\omega\rangle$ and $\sigma(\omega)$ represent respectively the average value and the standard deviation of the overlap computed on an ensemble of suitably randomised multiplex networks. In particular, for each continental airline system we sampled 1000 multi-layer graphs from the configuration model which maintains fixed the degree sequence of all the layers and rearranges the links on each layer, pairing edge stubs at random. We computed also $z(h_{\max})$ and $z(\eta_{\min})$, i.e. the z-scores of the maximal entropy rate and minimum variance over those 1000 multiplex graphs.

The results reported in Fig. 7.5 confirm that also in real-world systems $h_{\max}$ is negatively correlated with edge overlap, in agreement with the results obtained on synthetic networks. Similarly, $\eta_{\max}$ is positively correlated with ω. Notice that we have $z(\omega) > 0$ in all the six continents, meaning that the edge overlap of the real-world systems is always higher than that of the null-model, in agreement with the observation that real-world transportation networks tend to guarantee a certain level of robustness to failures. However, the quest for robustness has a cost in terms of dispersiveness and accessibility. In fact, $h_{\max}$ is consistently smaller than the value observed in the randomised systems ($z(h_{\max}) < 0$) for all continents, and similarly the steady-state probability distribution is consistently larger than that observed in the null model ($z(\eta_{\min}) > 0$)

It is quite interesting to note that the two multiplex networks with smallest overlap and overall better diffusion properties are the continental networks of Oceania and Europe, which span the least geographical space. We can speculate that in such systems some nodes representing cities in different countries are connected comparably well by different modes of transport, such as trains and bus, suited for relatively short distances and not included in our analysis. This might potentially explain the relative low number of redundant edges in those two airline transportation systems. Conversely, the necessity to provide route redundancy has somehow forced the airline transportation networks of Asia, South America and North America towards slightly less efficient configurations.

7.5 Discussion

In this Chapter we have explored how to extend biased random walks to the case of multiplex networks, showing that the richness of multi-layer systems allows to define several different classes of walks. In particular we studied the general features of the so-called

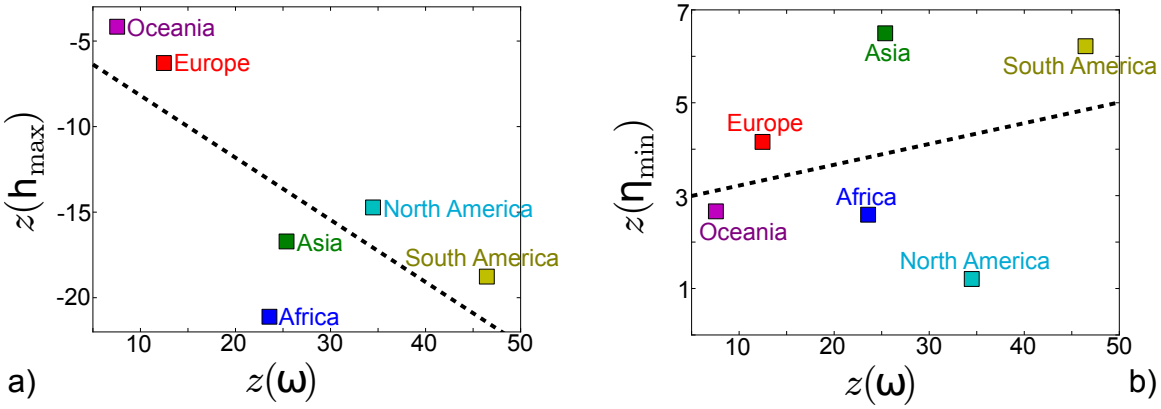


Figure 7.5: z-score of the average edge overlap ω versus the z-scores of the maximal entropy rate $h_{\max}$ (a) and the minimum standard deviation of the stationary distribution $\eta_{\min}$ (b) obtained through intensive walks. In agreement with findings for synthetic networks, also in real-world systems $z(h_{\max})$ is negatively correlated with $z(\omega)$ - Pearson's correlation coefficient $r = -0.70$ -, whereas $z(\eta_{\min})$ is positively correlated - $r = 0.30$, which increases to $r = 0.67$ excluding the outliner North America - [165].

extensive walkers (where the node properties, as the degree, are considered separately at each of the layers with different biasing parameters) and intensive walkers (biased on of the product two intrinsically multiplex, namely the overlapping degree and the participation coefficient) finding closed forms for the stationary occupation probability of these walks and for the entropy rate, and provided simplified heterogeneous mean-field expressions for the case in which the multiplex has no correlations.

We thoroughly investigated how structural properties of the multiplex, such as its number of layers, the presence of edge overlap and/or inter-layer degree correlations, the density of the layers and the heterogeneity of their degree distribution affect the dynamics of the random walkers. We found that number of layers, edge overlap and inter-layer degree correlations have a substantial impact on the diffusion properties of the walks. Also, we found that intensive random walkers perform at least as well as extensive random walkers in all the considered scenarios, with the advantage that the number of bias parameters does not scale with the number of layers.

Finally, the study of the diffusion properties of six real-world multiplex networks, namely the continental airline transportation networks of Africa, Asia, Europe, Oceania, North and South America, has shed some new light on the interplay between efficiency and robustness in multi-layer transportation systems. In particular, we found that the emerging necessity to provide some resilience to single link failures, which corresponds

to the introduction of some level of edge overlap, has shaped these systems in such a way that their navigability, in terms of entropy rate and heterogeneity of the node occupation probability, has somehow been sacrificed in favour of robustness. The results of the present work represent a valuable theoretical contribution to the development of efficient strategies to explore, search or navigate multiplex networks, and confirm the importance of appropriately taking into account the multiplexity of interactions when modelling intrinsically multi-dimensional systems.

Chapter 8

A model with interacting opinions

The increasing availability of data sets about social relationships, such as friendship, collaboration, competition, and opinion formation, has recently spurred a renewed interest for the basic mechanisms underpinning human dynamics [170]. Aside with the classical studies in social sciences and social network analysis [48, 117, 171], some interesting contributions to the understanding of social dynamics have lately come from statistical physics [172], which has brought in the field new tools and analytical methods to study systems consisting of many interacting agents. In such wider context, much effort has been devoted to the study of the dynamics responsible for opinion formation in populations of interacting agents, and in particular to a more in-depth understanding of the elementary mechanism allowing the emergence of global consensus and of the role of endogenous and exogenous driving forces, including social pressure and mass media. As a result of this investigation, a plethora of models of opinion formation have been proposed and studied [173–180].

Although the majority of those models originally made the simplifying assumption of considering homogeneous interaction patterns, i.e. regular lattices, the rise of network science provided the tools to overcome this limitation, featuring more realistic interaction patterns. More recently, also the role of mass media in the formation of global consensus has attracted a lot of interest [181–186].

An aspect of social relationships that has been mostly discarded in the study of the emergence of consensus is the fact that agents usually interact in a variety of different contexts, making the interaction pattern effectively multilayered and multi-faceted. As a matter of fact, the urge to maintain a certain level of coherence among opinions on different but related subjects might actually play a crucial role in determining the reaction of each agent to external pressure and in facilitating (or hindering) the emergence of

global consensus. Moreover, the balance between the internal tendency towards coherence and the necessity to adequately respond to social pressure is naturally dependent on each person's attitude, thus implying a certain level of heterogeneity. Some individuals may be more prone to align more closely to the opinions of their neighbors in each of the different contexts where they interact, putting little or no importance to the overall coherence of their profile. On the contrary, some other agents may indeed be more reluctant to change their opinion on a topic, in spite of being urged by other individuals or media, if such a change results in a contradiction with another of their opinions on a different but related subject.

In this Chapter we propose a model of opinion formation that takes into account *i)* the concurrent participation of agents to distinct yet connected interaction levels (representing discussion topics or social spheres), *ii)* the presence of social pressure and *iii)* the exogenous action of mass media [187]. This analysis can be naturally cast in the framework of multiplex networks, which has recently proven successful for a more realistic modeling of different social dynamics [188–190]. According to this framework, agents are represented by nodes connected by links of different nature, where links of the same kind belong to the same layer of the system. Each layer thus represents the interaction pattern of individuals discussing a given topic. Different layers are in general endowed with different topologies, to mimic multi-layer real-world social systems where distinct interaction patterns are present at different levels. Peer social pressure occurs on each topic through intra-layer links. The opinions of an individual on the different topics are also driven towards a specific state by the tension towards internal agent's coherence, represented by a preferred configuration of opinions on different topics. Mass media are introduced as fields acting uniformly on all the agents at the level of each single topic.

The resulting model is a natural extension of the traditional Ising model of magnetic interaction [191] and of more recent variations introduced to take into account the effect of external forces on the emergence of consensus [192], in the spirit of less and more recent work connecting statistical mechanics of disordered systems and opinion dynamics [193, 194]. The key ingredient of heterogeneous distributed couplings between opinions lead to interesting equilibrium states, where agents can remain fully coherent while a variable level of global consensus is attained, depending on the strength of the pressure exerted by mass media. This clearly resembles the dynamics observed in real societies, thereby supporting the relevance of our approach.

8.1 Interacting opinions

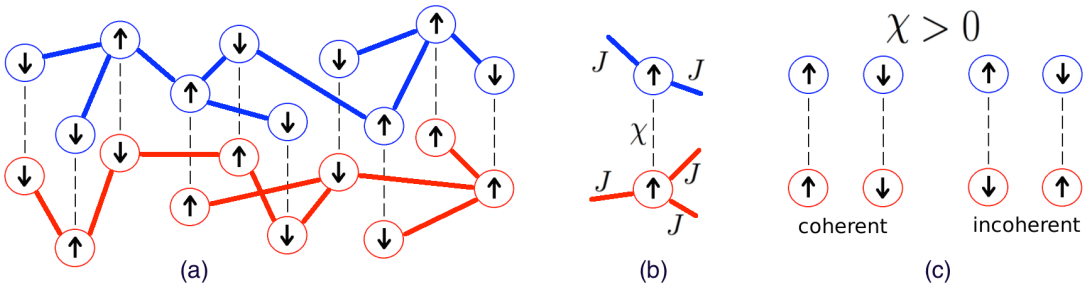


Figure 8.1: (a) Model of interacting opinions for $M = 2$ different topics. Every agent in the population expresses a binary opinion on each subject $\alpha = 1, 2$ (upward or downward arrows) and may interact with other individuals. The pattern of intra-layer interactions (blue/red lines for each α respectively) may in general be different. (b) Opinions may change according to both social pressure (weighted by the parameter J) and external fields, e.g. mass media (not shown). Individuals are also differently prone to change one or more of their opinions, according to their internal coherence. Such effect is taken into account by identifying natural couplings between the opinions of an agent at the two layers, weighted by the agent-dependent parameter χ_i. (c) The internal coherence/incoherence of an agent is determined by the sign of its opinions (orientation of the arrows): when $\chi_i > 0$, coherent configurations are those with both opinions of the same sign (left couple), whereas incoherent ones present opinions of different sign (right couple). The opposite holds if $\chi_i < 0$ [187].

We consider a population of N individuals interacting through M different layers, representing different topics or subjects. The network of each layer $\alpha = 1, \ldots, M$ represents the pattern of interactions among agents on a specific topic, which is in general distinct from those of the other layers. Each agent $i = 1, \ldots, N$ expresses a binary opinion $\sigma_i^{[\alpha]} = \pm 1$ on each subject $\alpha = 1, \ldots, M$. An example with $M = 2$ is shown in Fig. 8.1(a), where upwards and downwards spins represent the two possible values of $\sigma_i^{[\alpha]}$. We assume that agent opinions evolve over time due to two concurrent mechanisms. On the one hand, agents are subject to social pressure from their peers on each layer (denoted by the red and blue links in Fig. 8.1), so that the opinion of agent i on node α will tend to remain aligned with the opinions of its neighbors on the same layer. This mechanism, based on the elimination of conflicting opinions on a microscopic scale, has been widely observed in many real-world social systems, and is responsible for the attainment of local consensus on each layer. On the other hand, we assume that the opinions of agent i at the different layers are not independent from each other but

are instead interacting, so that for each agent there exists a preferred configuration of opinions at the different layers which is considered *coherent*. For instance, the political orientation of a person is often related to his/her ideas about economy and welfare, so that the emergence of consensus with its neighbors on one subject should remain coherent with its current opinions on the other layers. Moreover, we imagine that agents are exposed, on each layer, to the action of mass-media, a mean-field external force which preferentially drives their opinions towards either $+1$ or -1.

We formalize the interplay of these concurrent dynamics by defining the functional [187]:

$$f_i^{[\alpha]} = J \sum_{j=1}^{N} a_{ij}^{[\alpha]} \sigma_j^{[\alpha]} + h^{[\alpha]} + \gamma \frac{\chi_i}{J} \sum_{\beta=1,\neq\alpha}^{M} \sigma_i^{[\beta]} \tag{8.1}$$

for each agent i and each topic α. The first sum on the rhs of Eq. (8.1) represents the social pressure exerted on i by its neighbors on layer α, and is weighted by the coefficient J, which models its intrinsic permeability to social pressure. The variables $h^{[\alpha]}$ represent the external effect of mass-media on the formation of agents' opinions, which are considered in this case as a mean-field force acting homogeneously on all the agents of a layer. Finally, the second sum represents the tendency of agent i towards internal coherence, where the global parameter γ sets the relative importance of internal coherence and social pressure. Specifically, when $\gamma \simeq 0$ the opinions of the agents are mainly driven by peer and external pressure, whereas when $\gamma \to \infty$ they are determined by the internal coherence, such that coherent configurations are strongly favoured.

This setup is depicted in Fig. 8.1(b) for the case $M = 2$, where links with different colors indicate the connections of an agent at the two layers. In practice, J is the strength of the interaction of each agent with its neighbors, while χ_i determines the importance (and sign) of internal agent coherence. In this case, as shown in Fig. 8.1(c), the preferred configuration of agent's i spins is concordant if $\chi_i > 0$ and discordant if $\chi_i < 0$. We notice that the actual value of χ_i, which in the following always lie in the interval $[-1, 1]$, is a measure of how much agent i is flexible towards a change of one of its opinions, eventually leading to configurations which do not agree with what it would consider a coherent configuration of its spins. In other words, agents for which $|\chi_i| \simeq 0$ assign less importance to internal coherence and more relevance to social pressure, while the opposite happens when $|\chi_i| \simeq 1$.

In our model, the opinions of each agent evolve towards configurations which maximize the function $F_i^{[\alpha]} = \sigma_i^{[\alpha]} f_i^{[\alpha]}$, in order to attain, at the same time, internal coherence and local consensus with their neighbors on each layer [187]. As a consequence, agents will naturally prefer configurations of spins at all layers which ensure a balanced trade-off

between social pressure and coherence, depending on the respective values of the parameters γ, J, χ_i, and of the external fields $h^{[\alpha]}$. Although being a somehow simplified model of real-life interactions, where not just binary but also intermediate opinions between two extremes are possible and agents might respond differently to social pressure and to the external effect of mass-media, this model turns out to be already general enough to investigate the elementary mechanisms driving interacting opinions.

Numerical implementation of this dynamical evolution is obtained through extensive Monte Carlo simulations, adopting an appropriately modified version of the Glauber algorithm [195]. In particular, at each step we update all the spins $\sigma_i^{[\alpha]}$, $i = 1, \ldots, N$, $\alpha = 1, 2, \ldots, M$ in a random order. The update is performed by proposing a flip of the current spin $\sigma_i^{[\alpha]} \to -\sigma_i^{[\alpha]}$ and accepting the flip only when the new configuration leads to a larger value of the function $F_i^{[\alpha]}$. Every time a flip is accepted, $F_i^{[\alpha]}$ is also updated according to the new configuration. Clearly, the form of $f_i^{[\alpha]}$ captures both the contributions of intra-topic and inter-topic couplings and those of the existing external fields, so that larger values of $F_i^{[\alpha]}$ correspond to preferred configurations for node i.

Clearly, these rules imply a deterministic evolution of the opinions, which is not observed in real social systems. We then need to account for the presence of stochastic noise. Its simulation is realized by introducing a parameter $T \geq 0$, which may be regarded as a social temperature in analogy with magnetic systems, induced by all those mechanisms which drive the system out of its deterministic dynamics, such as partial information or misunderstandings. We include such thermal noise in the dynamics of our model in a standard way: when $T > 0$, an agent i may change its opinion on the topic α even if it leads to configurations with a smaller $F_i^{[\alpha]}$ with probability $e^{\frac{\Delta F_i^{[\alpha]}}{T}}$, with $\Delta F_i^{[\alpha]}$ being the variation in $F_i^{[\alpha]}$ due to the flip of thew spin $\sigma_i^{[\alpha]}$.

8.2 Consensus and coherence

We are interested in understanding how the presence of three concurrent factors, namely the response to social pressure, the tension towards internal coherence and the presence of an external mean-field force on each layer, affects the emergence of consensus in the population. We consider three scenarios, namely *i)* the case in which $\chi_i = 1, \forall i$ (homogeneous agents); *ii)* the case in which χ_i is a random variable sampled from a certain probability distribution (heterogeneous agents); and finally *iii)* the case in which the dynamics is affected by noise. For the sake of simplicity, in the following we focus on the case of two interconnected layers, i.e. $M = 2$, and we set $J = 1$, so that the relative importance of internal coherence and social pressure is determined, for each

agent, by the product $\gamma\chi_i$. The model is numerically investigated in the three different setups described above, where the parameters γ, $h^{[1]}$, and $h^{[2]}$ play the role of control parameters.

We study the emergence of consensus at each layer $\alpha = 1, 2$ through the order parameter [187]:

$$M^{[\alpha]} = \frac{1}{N}\sum_{i=1}^{N}\sigma_i^{[\alpha]}, \tag{8.2}$$

which satisfies $-1 \leq M^{[\alpha]} \leq 1$, with $\left|M^{[\alpha]}\right|$ denoting the strength of consensus and $\mathrm{sgnM}^{[\alpha]}$ indicating which of the opinion is prevalent among the population. We also define the average internal coherence [187] of the agents as follows:

$$C = \frac{1}{N}\sum_{i}^{N}\mathrm{sgn}(\chi_\mathrm{i})\sigma_\mathrm{i}^{[1]}\sigma_\mathrm{i}^{[2]}. \tag{8.3}$$

Notice that, when $\gamma > 0$, $C = +1$ if the two spins of each agent are coherent with their preferred configuration, while $C = -1$ if they are incoherent for every agent. The opposite holds when $\gamma < 0$.

An interesting remark is that the global function

$$H = -J\sum_{\alpha=1}^{M}\sum_{i,j=1}^{N}a_{ij}^{[\alpha]}\sigma_i^{[\alpha]}\sigma_j^{[\alpha]} + \sum_{\alpha=1}^{M}\sum_{i=1}^{N}h^{[\alpha]}\sigma_i^{[\alpha]} + \frac{\gamma}{J}\sum_{\alpha,\beta=1}^{M}\sum_{i=1}^{N}\chi_i\sigma_i^{[\alpha]}\sigma_i^{[\beta]} \tag{8.4}$$

(which we do not consider in this study) effectively is the Hamiltonian of a multi-layer Ising model, where the population evolves equivalently towards configurations that minimise H. In this sense, our model can be considered as a generalization of the coupled Ising model on lattices [196]. Following this analogy, we note that the order parameters $M^{[\alpha]}$ can be interpreted as the magnetization of the different layers of the system.

We discuss in the following sections the transition towards coherence and consensus and the equilibrium properties of the model, focusing on the dependence of the order parameters C and $M^{[\alpha]}$ in Eqs. (8.2-8.3) on the parameter γ and on the external fields $h^{[\alpha]}$ [187]. In details, we investigate in Sec. 8.3 the case of $\chi_i = 1\ \forall i$ and $T = 0$, i.e. a population of homogeneous agents in the absence of social noise. In Sec. 8.4 we consider a population of heterogeneous agents (χ_i not fixed), while keeping $T = 0$. Finally, in Sec. 8.5 we study the effect of social noise by investigating the dependence on T.

Simulations of the Glauber dynamics described in the previous section are realized by varying the global parameter γ adiabatically. The initial configuration is obtained by

setting $\sigma_i^{[1]} = 1$ and $\sigma_i^{[2]} = -1$ $\forall i$. We let the system perform two complete hysteresis cycles before recording the resulting configurations. This procedure eliminates possible effects due to the specific initial conditions.

The results presented here are obtained by simulating the dynamics on a multiplex of two random graphs with the same average degree $\langle k \rangle$=6.

8.3 Homogeneous agents: transition towards full coherence

We consider here the case of homogeneous agents $\chi_i = 1$ $\forall i$, in the absence of social noise, i.e., the case $T = 0$. The effects induced by the external forces, e.g., the mass media, are studied by choosing fields with opposite signs and relative strength according to the two typical cases: $|h^{[1]}| = |h^{[2]}|$ or $|h^{[1]}| > |h^{[2]}|$. We remark that the qualitative behaviour observed does not depend on the specific values of $|h^{[1]}|$ and $|h^{[2]}|$. First, we study the transition in coherence as a function of γ: for fields of both equal and different intensity, we provide evidence of the existence of a sharp transition along with a hysteresis loop. We are also able to propose an empirical relation to estimate the transition points $\gamma_{\pm}$, given the intensity of the fields and the density of the layers. We note that the case of fields with equal signs is somehow trivial, since the opinions on both layers are pulled in the same direction and global consensus emerges easily. Second, we find that a coherent population, i.e. in the regime $\gamma \to \infty$, exhibits either states of full or null consensus and that states of partial consensus cannot be attained in a population of homogenous agents [187].

We show examples of the steep transitions that the system exhibits by plotting C as a function of γ in the top panels of Fig. 8.2(a1) for $|h^{[1]}| > |h^{[2]}|$ and of Fig. 8.2(a2) for $|h^{[1]}| = |h^{[2]}|$. The behavior of the coherence is robust with respect to the relative strength of the external fields: we always observe a sharp transition from $C = -1$ to $C = +1$ characterised by a marked hysteresis loop. However, the actual values of $h^{[1]}$ and $h^{[2]}$ deeply affect the corresponding level of consensus emerging in the population. This is shown in the bottom panels of Figs. 8.2(a-b), where we plot the corresponding value of $M^{[2]}$ as a function of γ. If the external fields have the same intensity $|h^{[1]}| = |h^{[2]}|$, we have $M^{[2]} = 0$ for $\gamma > \gamma_+$, while $M^{[2]} = -1$ when $\gamma_- \leq \gamma \leq \gamma_+$. As the transition is sharp, we can always infer the value of $M^{[1]}$ from the corresponding values of C and $M^{[2]}$. In fact, we respectively have $M^{[1]} = \pm M^{[2]}$ when $C = \pm 1$.

This result has a clear interpretation. When γ is increased, the second term in

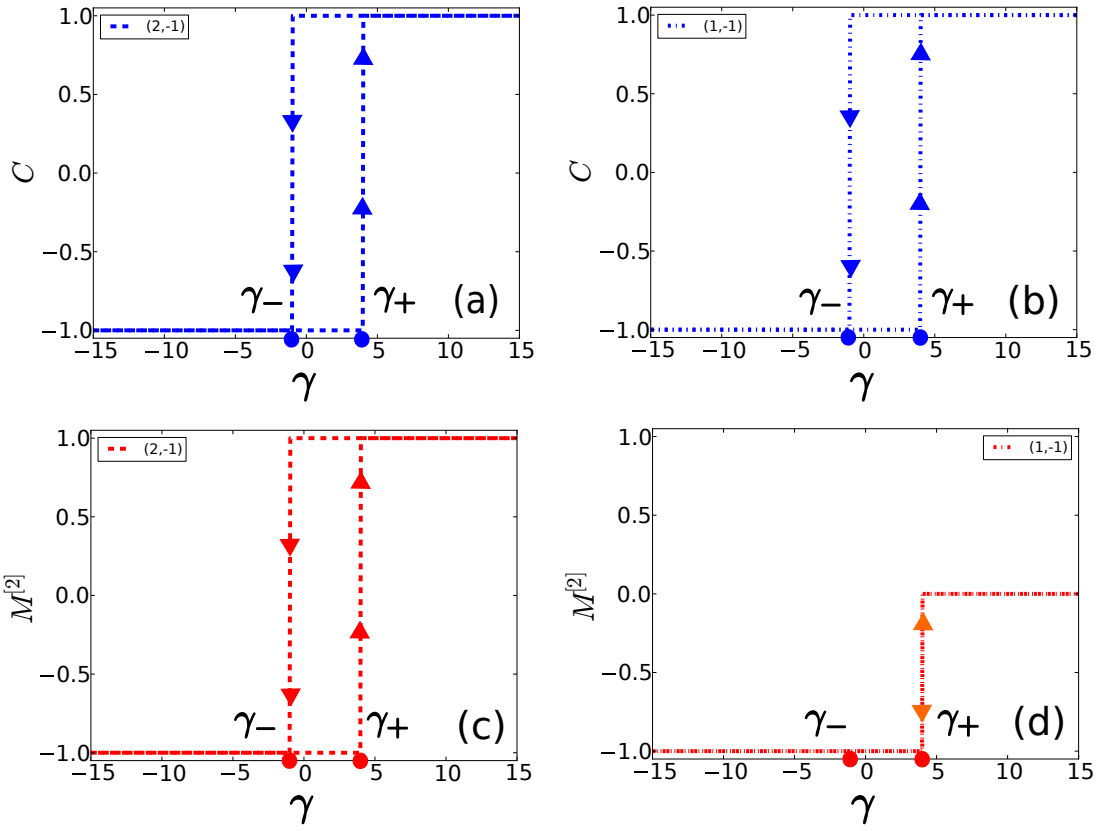


Figure 8.2: Values of the average agent coherence C (top panels) and of the consensus on the second topic $M^{[2]}$ (bottom panels) as a function of γ for a population of homogeneous agents ($\chi_i = 1$, $\forall i$). External fields are chosen with opposite signs and either different ($|h^{[1]}| > |h^{[2]}|$) (left panels) or equal ($|h^{[1]}| = |h^{[2]}|$) relative intensity (right panels). A sharp transition towards full coherence ($C = 1$), characterised by an hysteresis cycle delimited by the transition points γ_+ and γ_-, is observed in both cases (panels a-b), independently from the relative strength of the media. Conversely, the corresponding value of $M^{[2]}$ after the transition, i.e., when agents are coherent, differs significantly: if $|h^{[1]}| = |h^{[2]}|$ (panel d), we have $M^{[2]} = 0$, while if $|h^{[1]}| > |h^{[2]}|$ (panel c), we have $M^{[2]} = 1$. The presence of a stronger media pressure on a specific topic indeed influences also the other one. We note that when agents are homogeneous no states of partial consensus are allowed on either layer [187].

the right hand side of Eq. (8.1) becomes dominant, meaning that the agents give more importance to internal coherence than to social pressure. At the same time, however, none of the two external fields, which have opposite signs, is able to force a flip of opinions

on the the other layer. This leads naturally to states of vanishing magnetization on each layer, i.e., no consensus. The opposite situation is observed when we decrease γ. Indeed, the agents become more flexible, so that different opinions on different topics can coexist. Of course, this tendency gradually increases the effect of the external fields on their own topic. At the transition point, the population becomes globally incoherent, whereas the external fields induce full consensus separately on each layer, with their sign determining the dominant opinion.

In the case where one of the two fields is larger than the other, i.e. $|h^{[1]}| > |h^{[2]}|$, the situation is radically different. We indeed find $M^{[2]} = +1$ for $\gamma > \gamma_+$, $M^{[2]} = -1$ when γ increases in the interval $[\gamma_-, \gamma_+]$, and $M^{[2]} = +1$ when γ decreases in $[\gamma_-, \gamma_+]$. The interpretation follows straightforwardly with a reasoning similar to the one reported above for the case $|h^{[1]}| = |h^{[2]}|$. As γ increases, the agents become more and more inflexible, thus favoring opinions of the same sign throughout the different topics. Moreover, since $|h^{[1]}|$ is larger than $|h^{[2]}|$, states of non-vanishing consensus are favored. In particular, one of the opinions ends up prevailing not just on layer 1 but, through the internal agent coherence, also on the other layer. Thus, the concurrent effect of these two mechanisms causes a steep transition towards a state of both full coherence and full consensus on a single opinion on both the topics, which is determined by the leading external field. The same dynamical explanation of the previous case can instead be given for decreasing values of γ beyond γ_-.

As suggested before, these qualitative patterns are robust with respect to the strengths of the external fields, which only determine the exact transition points γ_+ and γ_-, as shown in Fig. 8.3(a). We find that the transitions points γ_+ and γ_- where the hysteresis loop starts and ends respectively are given by the following empirical non-linear relation:

$$\gamma_\pm = \pm\langle k\rangle/2 - \mathrm{sgn}(\mathrm{h}^{[1]}\mathrm{h}^{[2]})\,\min(|\mathrm{h}^{[1]}|, |\mathrm{h}^{[2]}|), \tag{8.5}$$

where $\langle k\rangle = \frac{1}{2N}\sum_{\alpha=1}^{2}\sum_{i,j=1}^{N} a_{ij}^{[\alpha]}$ is the average degree of the two layers. The actual values of $h^{[1]}, h^{[2]}$ only determine a shift of the metastable region, whereas they do not modify the width of the hysteresis cycle. We support this conjecture by showing in Fig. 8.3(b) the values of $(\gamma_+ + \gamma_-)/2$ (i.e. the center of the hysteresis cycle) obtained from the simulations as a function of $h^{[1]}$ and $h^{[2]}$, confirming the validity of the relation expressed in Eq. (8.5).

We conclude that in the case of homogeneous agents the system always reaches configurations of full consensus on both layers, where the dominant opinion on each layer is determined by the sign of the strongest external field (phase diagram in Fig. 8.4, top panel a). The only exception is given by the critical line $h_1 = -h_2$ where we find $M^{[2]} \approx 0$.

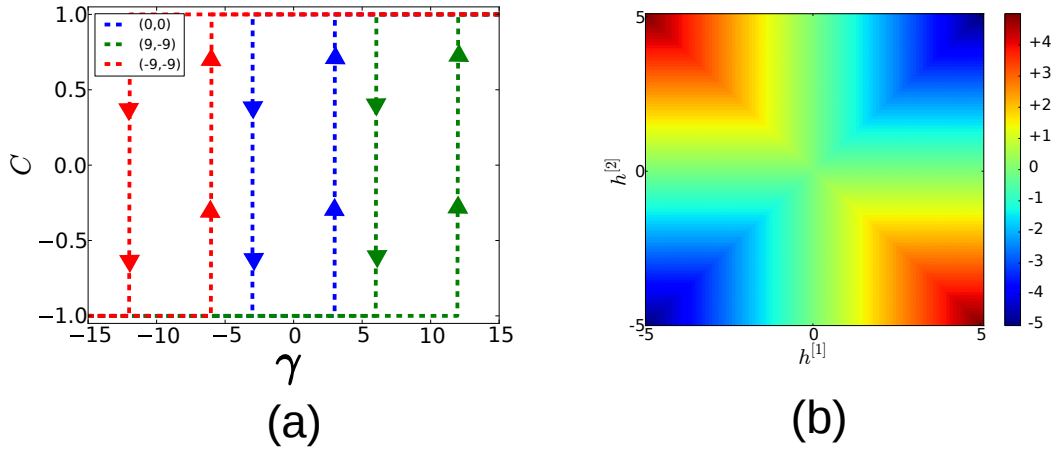


Figure 8.3: (a) Plot of the hysteresis cycle of the average internal coherence C for different values of their external fields. Even if the qualitative behaviour of the system does not depend on the actual values of $h^{[1]}$ and $h^{[2]}$) (i.e., the system is characterised by a sharp transition in C with a marked hysteresis loop, whose width is determined solely by the average degree on the two layers $\langle k \rangle$), the exact positions of the transition points γ_+ and γ_- change according to the intensity and sign of the external fields. (b) Simulated values of $(\gamma_+ + \gamma_-)/2$ as a function of $h^{[1]}$ and $h^{[2]}$. These numerical results support the validity of the empirical relation of Eq. (8.5).

As expected, the assumption of homogeneity of the agents imposes a strong constraint on the dynamics of the model, leading only to unrealistic patterns of perfect (or null in the specified particular case) consensus always accompanied by perfect coherence, but not allowing intermediate configurations. These sharp scenarios are different from those observed for real-world systems, where states of partial consensus are often observed. The heterogeneity in the relative weight and sign assigned to internal coherence indeed plays a crucial role by favoring the influence of the mass media over the attainment of full coherence in the population. Thus, we expect that the relaxation of the homogeneity hypothesis in our model could lead to milder patterns, with different levels of consensus at equilibrium, thus better resembling the observed dynamics of real-world societies.

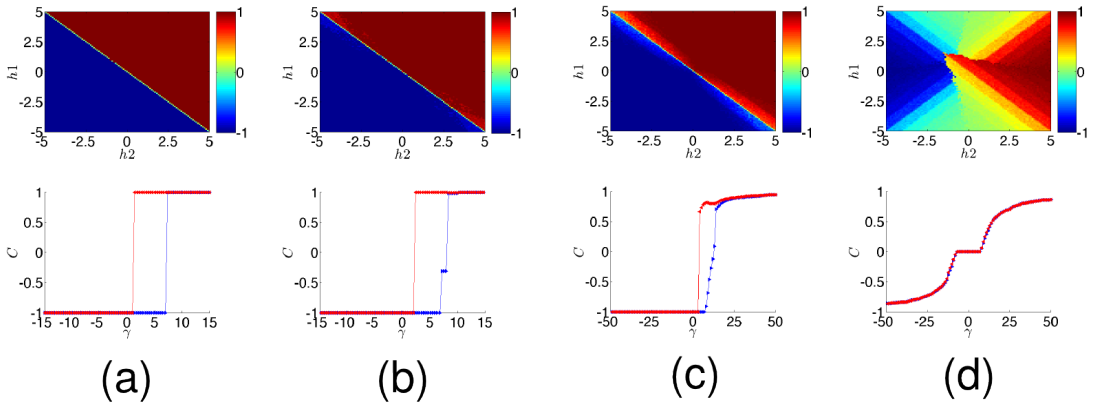


Figure 8.4: Phase diagrams of $M^{[2]}$ (consensus on the second layer) with respect to the external fields (top panels) for a population of coherent agents, i.e., for γ large enough to have average coherence $C = 1$ ($\gamma = 15$ for panels (a-b) and $\gamma = 50$ for panels (c-d)), and the corresponding transition of C (bottom panels) as a function of γ (forward/backward branch of the hysteresis cycle respectively for blue/red lines) for $h^{[1]} = 5$, $h^{[2]} = -3$. The inter-layer coupling χ_i for each panel is sampled from different distributions, namely (a) $\chi_1 = 1\ \forall i$, (b) half of the agents with $\chi_i = 1$ and the remaining half with $\chi_i = 1/2$, (c) χ_i uniformly distributed in $[0, 1]$, and (d) χ_i uniformly distributed in $[-1, 1]$. We recall that in this regime, where $C = 1$, we have $M^{[1]} = M^{[2]}$ [187].

8.4 Heterogeneous agents: emergence of partial consensus states

We consider here the case of a population of heterogeneous agents, i.e. χ_i may be different for each agent i [187]. As in the previous section, we set $T = 0$, meaning that we neglect the effect of social noise. Such realistic scenarios break the steep transition of the average internal coherence C and allow for the emergence of states of partial consensus in populations of coherent agents. We support this claim by reporting in Fig. 8.4 both the phase diagrams of the consensus on the second layer $M^{[2]}$ for $\gamma \gg 1$, or equivalently $C = 1$, (top panels) and the plot of C as a function of γ for a typical choice of the external fields ($h^{[1]} = 5$, $h^{[2]} = -3$ specifically) for a few simple but explanatory cases.

We first consider the simplest possible setup where half of the population is assigned $\chi_i = 1/2$, whereas the other one is assigned $\chi_i = 1$ [Fig. 8.4(b)], meaning respectively that 50% of the population is flexible with respect to internal consensus ($\chi_i = 1/2$) while

the remaining agents are intransigent ($\chi_i = 1$). Even if in this case the phase diagram looks similar to the one in Fig. 8.4(a) for a population of homogeneous agents, we can already observe the emergence of states of partial consensus close to the diagonal, i.e., for $|h^{[1]}|, |h^{[2]}| > 2.5$. The breaking of the steep transition in C is also confirmed in the bottom panel of Fig. 8.4(b).

We then consider in Fig. 8.4(c) the case of an heterogeneous population with $\chi_i \in U(0, 1)$, i.e., uniformly distributed in the interval $[0, 1]$. In this case, the qualitative behaviour of both $M^{[2]}$ and C is similar to the one shown in Fig. 8.4(b). However, as expected due to the increase of the level of heterogeneity of the population, the regions of partial consensus are wider and characterized by lower values of $M^{[2]}$ with respect to the previous case.

Thus, we may expect to find even richer phase diagrams and smoother transitions in C with respect to the cases presented before if we further increase the heterogeneity of the population. Indeed, when χ_i is sampled uniformly in $[-1, 1]$, the phase diagram looks qualitatively different: $M^{[2]}$ smoothly increases from -1 to $+1$ for increasing values of $h^{[2]}$ and fixed $h^{[1]}$. Furthermore, the consensus attained in the region $|h^{[2]}| < 2.5$ with $|h^{[1]}| > 2.5$ is significantly smaller than in the other cases. These results suggest that one can smoothly tune the level of consensus on each topic by choosing the relative strength of the media acting on the two layers, and yet obtain states in which the majority of the agents are internally coherent. We also recall that in all the non-homogeneous cases (Fig. 8.4(b-d)) the system reaches full coherence, but the transition is not sharp. We conclude by highlighting that our model, even if simplified, is nevertheless able to generate non trivial states of partial consensus across the layers due to the driving effect of mass media, while at the same time ensuring that each agent will still find itself coherent.

8.5 The effect of noise

We here consider the case with social noise, i.e. $T > 0$. For simplicity, we investigate its effect in a population of homogeneous agents ($\chi_i = 1\ \forall i$). We find that the system exhibits the same qualitative behaviour described in the case $T = 0$ for all temperatures below a non-null critical temperature T_c, whereas for $T \geq T_c$ it does have absorbing states for finite values of γ, thus lying in a paramagnetic phase dominated by noise [187].

This is shown in Fig. 8.5(a) where we plot C as a function of both T and γ (forward branch of the hysteresis cycle) for an exemplary choice of the external fields $h^{[1]}$, $h^{[2]}$ with

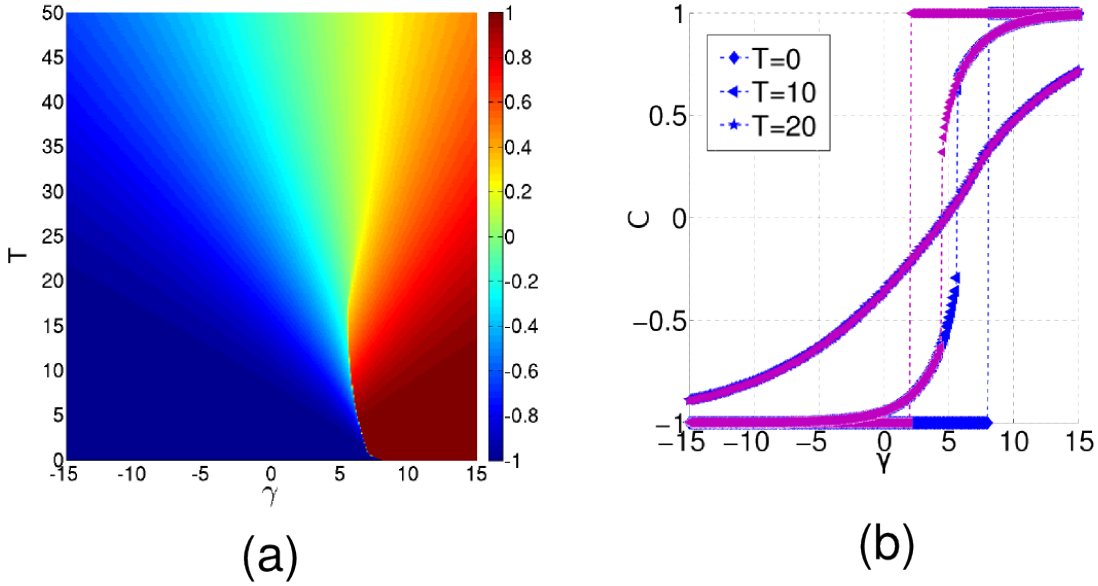


Figure 8.5: (a) Average internal coherence C as a function of the noise T and of the global parameter γ for $h^{[1]} = 10$, $h^{[2]} = -5$. Simulations are obtained by increasing γ adiabatically for fixed T (forward branches shown). There exists a critical value T_c of T below which the system always attains full coherence ($C = 1$) and consensus ($M^{[1]} = M^{[2]} = +1$) through sharp transitions (consensus not shown). For $T > T_c$ instead the noise becomes dominant and the agents remain incoherent for any γ finite. Eventually for $\gamma \to \infty$ full coherence is obtained via a smooth transition. (b) Projections of C for both increasing and decreasing values of γ (dark/light color respectively) for three different values of T (symbols). The transition between coherent and incoherent states smoothens as T increases, eventually becoming continuous when $T > T_c \approx 10$ [187].

opposite signs. We note that different values of $|h^{[1]}|$, $|h^{[2]}|$ do not change qualitatively the results presented. Indeed, for $T < T_c$ the system exhibits steep transitions to states of full coherence and consensus. However, when T increases, i.e. as the noise becomes stronger, the jump of the transition becomes less pronounced and the hysteresis cycle shrinks considerably, eventually disappearing at $T = T_c$. For $T > T_c$ only states of partial coherence and consensus can be obtained, and $|C| \simeq 0$ for $T \gg T_c$. Only in the limit $\gamma \to \infty$, the population is able to recover full coherence.

This scenario is confirmed by Fig. 8.5(b), where we report projections of the phase diagram of Fig. 8.5(a) for different exemplary values of T. For $T = 0$ the hysteresis

cycle is wide and the jump in C goes from -1 to 1. For $T = 10$, slightly below T_c, the hysteresis cycle has almost disappeared and the jump in C is consistently reduced, though still present. For $T = 20$, i.e. beyond the critical level of noise, the transition in C becomes continuous and the hysteresis loop disappears. We note that the noise similarly affects the system in the case of a population of heterogeneous agents, such that a paramagnetic phase appears beyond T_c also in this case. Furthermore, we stress that T_c depends non trivially on the set of parameters of the system. However, deriving such functional relation is beyond the scope of the present work.

We conclude by recalling that opinion evolution in real social systems is often affected by noise. In this Section, we have shown that the behavior of the system for $T = 0$ does not change qualitatively in the presence of noise below some critical value T_c for both a population of homogeneous agents and one of heterogeneous agents. This ultimately suggests that our finding that heterogeneity is necessary in population of coherent agents in order to exhibit realistic states of partial consensus, found for noise-free setups of our model, may still be relevant for real social systems.

8.6 Discussion

Understanding the elementary mechanisms responsible for the emergence of consensus in social systems is a fascinating problem that has stimulated research in several different fields, from sociology to mathematics, from computer science to theoretical physics, for more than a few decades. Nevertheless, traditional models used in the field to describe such systems are still far from capturing the essence of the dynamics of real societies.

Indeed, these models of opinion formation overall underestimate the importance of both (i) the existence of many different contexts where social dynamics may develop, and (ii) the variety of interaction patterns that naturally forms between individuals at each of these different aspects. In details, these models are usually based on the simplifying assumption that the social interactions underpinning consensus are essentially homogeneous, whereas real-world societies are instead intrinsically multilayered and multifaceted, meaning that individuals normally interact with several different neighbourhoods in a number of different yet correlated contexts. Such multilayered structure of social interactions also naturally imply that relationships among each individuals' opinions on many different topics or subjects may exists, thus playing a major role in the formation of an agent's public profile. However, this issue has rarely been addressed in the literature to our knowledge. Overall, these properties of real social systems, force agents to pursue a balanced trade-off between their internal tendency towards providing a

coherent image of themselves, corresponding to a coherent set of opinions over the range of contexts in which their social activities develop, and the external pressure towards local homogeneization that comes from their concurrent participation to different social circles.

In this Chapter we address the issues (i-ii) thoroughly, and propose a novel, yet simple, model of social opinion dynamics which is capable to account for them all. Our model is obtained by suitably readapting the framework of multilayer networks, which has been developed in the last years in different contexts. Remarkably, the proposed model suggests that the delicate equilibrium between internal agent coherence and responsiveness to external social pressure in a multilayered social environment might indeed be one of the fundamental ingredients responsible for the appearance of non-trivial consensus patterns, such as states of partial consensus emerging from a population of coherent agents. Despite being straightforward in its formulation and relying on rather simple assumptions, the model we proposed allows to take appropriately into account the interplay between each agent's tendency towards coherence, the neighborhood's tendency towards local consensus and the pulling external forces represented by the persistent action of mass media. One of the most interesting findings of the present work is that the introduction of mild heterogeneity in the agents' response to social pressure fosters the emergence of non-trivial states in which internal agent's coherence is always reached at the expenses of a lower level of global consensus. This picture is consistent with what is widely observed in structured societies [197], where a perfect global consensus is never stable while individuals tend to adhere to pre-defined sets of social values which they consider coherent.

Another remarkable effect reproduced by our model is the impact of mass media pressure, especially in the case where the population is heterogeneous. In particular, it is interesting to observe that by an appropriate tuning of the relative strength of the two external fields representing mass media one can indeed set any desired value of consensus on each layer, with the possibility of driving the population from incoherent to more coherent configurations in a continuous way. Finally, the results of the study of the role played by the presence of noise are compatible with real-world scenarios, in which incomplete or inaccurate information about the state of peers is the norm and not an exception.

We highlight that the model discussed in this work is limited to a specific setting, where both the social and mass-media pressure are considered only as a mean-field effect. These assumptions imply that the response of agents to both external fields and interactions with his/her neighbors is homogeneous, which is only a first-order approximation of the real effects of mass media and social pressure on a population of agents.

A more realistic approach would require to consider each agent's adaptive response to such influence, i.e., by both considering that the effect of external field on layer α on each node i is a random variable $h_i^{[\alpha]}$ drawn from a certain distribution, and considering an agent-dependent response to interactions with other individuals, i.e. by replacing J with an agent-dependent parameter J_i. However, we purposedly decided to leave the investigation of these generalizations to a future work.

In conclusion, we find it quite intriguing that by taking into account the presence of concurrent interactions on a variety of different topics we were able to provide a simple explanation for the formation of growing patterns of consensus, whose level appears to be dependent on the strength of mass media pressure, as long as the agents acknowledge different couplings between their opinions on the different topics. We believe that the results presented in this work will spur further research towards a better understanding of the implications of interconnected and multilayered interaction patterns on the spreading of opinions and emergence of consensus in real-world social systems.

Chapter 9

Cultural dynamics: revisiting the Axelrod model

Despite the presence of increasing pressure towards globalisation, the coexistence of different cultures is a distinctive feature of human societies. However, how multiculturality can emerge in a population of individuals inclined to imitation, and how it remains stable under cultural drift, i.e. the spontaneous mutation of traits in the population, still needs to be understood. To solve such a problem, we propose in this Chapter a microscopic model of culture dissemination which takes into account that, in real social systems, the interactions are organised in various layers corresponding to different interests or topics. We show that the addition of multiplexity in the modeling of our society generates qualitatively novel dynamical behavior, producing a new stable regime of cultural diversity [198]. This finding suggests that the layered organisation of social influence typical of modern societies is the key ingredient to explain why and how multiculturality emerges and thrives in our world.

9.1 Imitation, drift and multiculturality: a difficult puzzle

The existence of multiculturality and group boundaries is a well-established feature of social systems [199, 200] and much effort has been devoted into explaining possible mechanisms able to reproduce such empirical finding. In particular, cultural diversity appears as a striking phenomenon, supposedly in contrast with the widely acknowledged principle of social influence [201]. According to such mechanism, changes in the cultural traits of an agent are influenced by the acquaintances and friends of that agent facilitating the local convergence towards a set of common cultural traits and promoting homogeneity

across the populations. For instance, homogenous absorbing states are typical of imitative dynamics where the state of the individuals is described through a scalar binary variable, as for the voter model [202] in finite size populations.

An interesting solution to reconcile social influence and the emergence of cultural diversity commonly observed in human societies was suggested by Robert Axelrod, who proposed in 1997 a simple agent-based mechanistic model of dissemination of cultural traits [179]. In the original model the state of individuals is not described through a scalar variable, but each agent is endowed with a set of cultural features F, each of them taking one of a number of different cultural traits q. Based on pairwise interactions among agents, together with social influence the model mimics another important social principle, known as homophily [203–205], i.e., the tendency of individuals to connect and interact preferentially with similar ones. As a result of pairwise interactions, imitation still occurs but is limited to the update of one feature at a time. Homophily and social influence acting together constitute a self-reinforcing dynamics leading to local homogenization. However, a main result of the analysis of Axelrod was that in spite of local convergence, global polarization was possible: agents become more similar by local interactions, but cleavages among different cultural groups are created so that these groups no longer interact. The overall result is the possible emergence of a globally polarized or multicultural state with coexistence of different cultural groups. A quantitative analysis of the model [206] unveiled a non-equilibrium phase transition at a critical number of cultural traits q_c from a mono-cultural phase (global culture) to a polarized or multicultural phase, where several groups with different cultural traits survive. In the previous chapter we investigated what happens to a simple spin-model, historically at the heart of statistical physics and opinion dynamics, when two systems are coupled together. In this chapter we revisit the famous social dynamics proposed by Axelrod with tools and ideas from multiplex network theory. We note that, differently from other social dynamics, also in its simplest implementation on single-layer networks the Axelrod model is somehow inherently layered, as the state of each individual is described by a set of different spins associated to different traits. For such a reason the Axelrod model and all its modifications are not considered to be models of opinion dynamics, but rather to form a class of their own, usually referred to as cultural dynamics. The model was at first implemented on regular lattices but the same transition from globalisation to multiculturality has been observed also for realistic interaction patterns, such as small-world networks and heterogeneous distributions of the connections among the individuals [207].

However, it was found that the multicultural phase is achievable only for a high initial number of traits q in the population. In addition, the multicultural phase is not stable under cultural drift, meaning that the spontaneous tendency of agents to modify their

cultural traits independently of their environment drives the system towards a monocultural state [208, 209]. Different mechanisms have been invoked to account for robust cultural diversity, such as the plasticity of the social relations among agents, so that the social networks coevolves with the dynamics of the cultural state [210]. Other alternative proposals focus on modifications of the form of the local interactions, for instance based on assimilation-contrast theory [211], also known as bounded confidence [212–214]. It has been shown how by integrating metric cultural states [215–217], as opposed to nominal states, the bounded confidence mechanism can lead to robust cultural diversity [218]. Another type of modification [219] addresses the nature of social influence in line with other models of social contagion [220–223]: the dyadic interpersonal influence of the original model is replaced by a mechanism in which an agent modifies its cultural state depending on the state of all agents in her neighborhood. Finally, the mechanism of social differentiation in which agents tend to increase cultural differences has also been analyzed in [224, 225].

In this Chapter we propose an alternative mechanism to account for robust multiculturality that naturally brings together social influence and the nature of the social network of interactions with no need of introducing any additional dynamical feature [198]. In real societies the relationships among individuals are inherently layered, i.e. individuals tend to interact with different neighborhoods on different topics. We model cultural dissemination associating each cultural feature to a different layer of a multiplex network. Social influence becomes now a layered mechanism in which an agent is allowed to imitate only a subset of all the cultural features of its neighbors, namely those for which there exists a link on the corresponding interaction layer. The consequences of layered social influence are strongly dependent on the structural overlap among different layers, so that the overlap becomes the control parameter for the nonequilibrium transition in the system.

We find that layered social influence in synthetic and empirical multiplex networks with heterogeneous layers easily leads to a global state of cultural diversity. This state exists for any number of cultural traits provided that the interaction patterns into the population are sufficiently layered, i.e. the value of the structural overlap is below a critical point [198]. We remark that this is a qualitative shift in the behavior of the system, only achievable in multiplex networks, and it is in agreement with empirical evidence of fragmented states even in social contexts with a limited number of available cultural choices q. Moreover, unlike the fragmented states found by Axelrod, such new multicultural state is robust against cultural drift, thus providing an explanation for the persistence of multiculturality we experience in real-world society. In addition, we find novel phases of cultural diversity in which a global culture for a number of cultural fea-

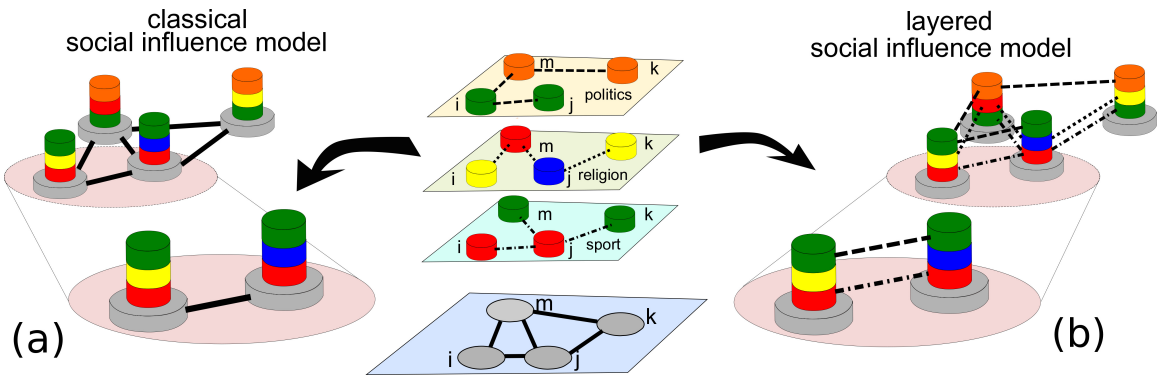


Figure 9.1: **Social influence is inherently layered.** Social systems are multiplex networks, in which each layer describes a different type of relations or interactions. Usually, the individuals tend to differentiate their social patterns, choosing recipients according to the subject of the interactions and viceversa. As an example, the couple of agents i and j discusses sports and politics, the couple j and k discusses sports and religion, while m and k can only discuss politics. In classical models of social influence such as the Axelrod model, the cultural profile of each individual is a vector of different features (sport, religion, politics), shown as a stack of cylinders, and each feature can take one of a series of different traits, represented as the colors of the cylinders. However, the peculiar nature of each interaction is neglected and all connections are treated together. Consequently, interactions at any level can potentially impact (and thus modify) any cultural trait (a). For instance, the existence of a link between nodes i and j, and the different traits of the pair on the feature "religion", suggest that this bond is still active. According to social influence, one of the two agents will eventually absorb the religious trait of the other one, making the pair equal on all the three features. Conversely, we propose a layered cultural influence model where the structure of the relationships pertaining to different social spheres is preserved and explicitly taken into account, assuming that each layered interaction can affect the corresponding cultural features (b). Agents i and j are not linked at the religious layer. As a consequence, since the two individuals already have the same sport and politics traits, the interaction between i and j is frozen. Indeed, the two individuals already reached consensus at all levels where they are linked, and imitation can not concern religious beliefs, since religion is not discussed by them [198].

tures coexist with polarization in other features, a situation reminiscent of the so-called chimera states [226–228]. More in general, we observe that different levels of homogeneity may be achieved on different cultural features, depending on the heterogeneity

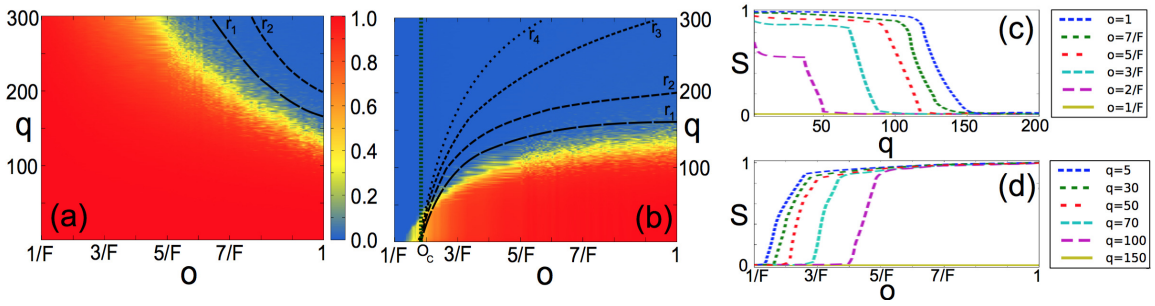


Figure 9.2: **Modeling classical versus layered social influence.** The size of the largest cultural component S in the classical (a) and in the layered (b) social influence model is shown as function of the number of cultural traits q, and for multiplex networks with $N = 625$ nodes, $F = 10$ layers, $\langle k \rangle = 4$, and a tuneable structural overlap o. In the case of classical social influence, a decrease in o increase the density of the aggregated networks, hence increasing the value of $q_c(o)$. Conversely, for layered social influence, the critical number of cultural traits $q_c(o)$ separating globalisation (red region) from multiculturality (blue region) decreases with the overlap, and goes to 0 for o smaller than a critical value. This means that, for sufficiently small values of the overlap, multiculturality is always achieved independently from the values of q. Under the presence of drift r, the critical overlap sustaining globalisation depends both on q and r. Noticeably, the polarised phase obtained for $o < o_c(q, r = 0) \approx 2/F$ is stable under the presence of noise. This is shown in (b), where we plot with different dashes the critical lines relative to four increasing values of noise $r = r_1, \ldots, r_4$. Conversely, the multicultural states on the classical model are not stable in presence of drift. The two projections of the phase diagram for layered social influence are reported in (c) and (d). Transitions from globalisation to multiculturality are steep both as a function of the number of cultural traits and the overlap [198].

of the structure of interactions across the layers. Finally, we investigate two layered social networks from the real-world, showing how considering the nature of the different interactions promotes cultural fragmentation in the population.

9.2 Cultural dissemination and layered social influence

We propose to take into account the multilayer structure of real social interactions by modelling layered social influence between pairs of individuals, as shown in Fig. 9.1. We describe the specific pattern of social influence at each cultural level f by mean of a

multilayer network described as the usual vector of adjacency matrix. Since each cultural feature is associated to a layer, in this Chapter we indicate the layers of a multiplex with the letter f rather than the usual α.

We consider once again the edge overlap, $o_{ij} = \frac{1}{F}\sum_{f=1}^{F} a_{ij}^{[f]}$ between two individuals, and each average value o among connected individuals, as introduced in Sec.2.3. Intuitively, when $o \approx 1$ the layered nature of social influence will be negligible. Indeed, if the peculiar nature of the links is neglected, the corresponding aggregated network of social interactions (formally defined at the end of this Section) will be similar to the interaction networks relative to each feature. Conversely, when $o \approx 1/F$ interactions are extremely diversified across the different topics. As a consequence, the aggregated network will be very different from the interactions occurring at each layer, hinting at possible different outcomes between the layered model and the original model on the aggregated network, to which we refer as the classical case.

The cultural profile of each agent is described by a feature vector of F integer variables $(\sigma^{[1]}, \ldots, \sigma^{[F]})$. Each feature f, with $f = 1, \ldots, F$, takes one of q possible traits, $\sigma^{[f]} = 1, 2 \ldots, q$. Because of the layered structure of interactions, when an agent interacts with a neighbor it only considers the subset of features where the two are connected. According to the principle of *homophily*, the probability of a social interaction between two connected agents i and j is assumed to be proportional to their layered *cultural overlap*, which we define as:

$$\omega_{ij} = \frac{1}{F}\sum_{f=1}^{F} a_{ij}^{[f]} \delta\Big(\sigma_i^{[f]}, \sigma_j^{[f]}\Big). \tag{9.1}$$

where $\delta\Big(\sigma_i^{[f]}, \sigma_j^{[f]}\Big)$ is the Kronecker's delta function. Notice that the cultural overlap of two agents is proportional to the number of shared traits on the layers in which the two agents are directly connected. When two agents interact, their cultural profiles are updated according to *layered social influence*. In practice, when agent i interacts with one of its neighbors j, imitation occurs, and j aligns one of its cultural features to that of i, choosing only among cultural features in which i and j interact, i.e., only those features for which a link between i and j exist in the corresponding layer. To avoid confusion between o and ω, in the following we will explicitly refer to the average edge overlap as to the *structural overlap*.

The dynamics of the model proceeds in epochs. During each epoch we consider each of the N agents in a random order, and perform the following steps:

1. Consider agent i, chosen uniformly at random.

2. Consider a neighbor j of i, chosen uniformly at random among all the neighbours of i (nodes with a link to i on at least one layer).

3. Let i and j interact with probability ω_{ij} in Eq. 9.1.

4. If the interaction takes place, choose at random a feature f such that $a_{ij}^{[f]} = 1$ and set $\sigma_j^{[f]} = \sigma_i^{[f]}$.

The dynamics stops when all pairs of connected agents i and j have either $\omega_{ij} = 0$ or $\omega_{ij} = o_{ij}$. This implies the existence and stability of interacting pairs of agents who only share a limited number of cultural features, a realistic property observed in real systems but not achievable in the classical model.

In fact, if we neglect the layered structure of interactions, we are implicitly assuming that any link between two individuals potentially allows for an exchange of traits over any feature. In that case, the structure of interactions in the systems can be described by the corresponding aggregated network $\mathcal{A}$. We note that the aggregate network does not lead to any discrepancy with the structure of the original multilayer networks for the dynamics of cultural diffusion when $o = 1$.

9.3 Emergence of stable multicultural states

In this Section we show how the presence of layered interactions, which mimic the multilayer structure of real-world societies, can explain empirical observations on the presence of multiculturality in real systems. In order to study the impact of layered interactions on culture diffusion, we model the structure of a social network by tuning its level of structural overlap o. We considered multilayer networks with F layers, each of them being an Erdös-Renyi graph with the same number of links such that $\langle k^{[f]} \rangle = \frac{1}{2N} \sum_{i,j \neq i} a_{ij}^{[f]} = 4$ $\forall f = 1, \ldots, F$.

In the case of F identical layers, the stuctural overlap is maximum and we have $o = 1$. Starting with this configuration, we then assign to each edge of each layer a probability p to be rewired at random. Since each edge is rewired independently, we can express the structural overlap o as a function of p. In particular, let us consider the case where each layer f has the same number of links $K^{[f]}$, so that we have $K^{[f]}F$ links in total. We then consider the aggregated network obtained by collapsing all the layers. If all the layers are identical, the aggregated graph will have $K = K^{[f]}$ edges. Hence $o = \frac{1}{F}\frac{KF}{K} = 1$ for $p = 0$. When $p \neq 0$ and links are rewired independently at each layer, the number of edges in the aggregated networks increases. If we neglect the probability that two links

of different layers are rewired in the same position, the number of edges in the aggregated network becomes equal to

$$K = \sum_{m=1}^{F} \binom{F}{m} (1-p)^m p^{F-m} K^{[f]} + \sum_{m=0}^{F} \binom{F}{m} (1-p)^m p^{F-m} K^{[f]} (F-m) \tag{9.2}$$

However the total number of connections in the multiplex remains $K^{[f]}F$. Hence

$$o \approx \frac{1}{\sum_{m=0}^{F} \binom{F}{m} (1-p)^m p^{F-m} (1 - \delta_{0,m} + F - m)}, \tag{9.3}$$

higher values of rewiring correspond to lower values of overlap and vice versa, with the limiting cases of $o = 1$ for $p = 0$ and $o = 1/F$ for $p = 1$. Given the average degree $\langle k^{[f]} \rangle$, the average number of neighbours $\langle k \rangle$ in the aggregated network can be obtained as $\langle k \rangle = \langle k^{[f]} \rangle / o$. Hence, the higher the p, the lower the o and the higher the density of the aggregated network.

In Fig. 9.2 we plot the size of the normalised largest cultural component S as a function of the number of cultural traits q and for different values of the structural overlap o, both for the classical Axelrod model (a), which runs on the aggregate graph associated to the multiplex network, and for the layered social influence model (b). Two individuals are considered to be part of the same cultural component if there is a path from one to the other (paths whose edges lay on different layers are allowed) and if they have equal cultural traits for all F features. In the classical model, a lower value of overlap by making the aggregated network denser is favouring globalisation, i.e. $S \approx 1$, increasing the critical number of cultural traits $q_c(o)$ at which multiculturality appears, i.e. $S \approx 0$.

For $o = 1$ (i.e., when all the layers are identical) the classical and layered model are undistinguishable, and multiculturality can only be achieved for large values of q. For instance, in the population considered in Fig. 9.2, multiculturality can be achieved only for $q > 140$. In the layered social influence model, where an agent can absorb the cultural trait of his neighbor only if they are connected instead, $q_c(o)$ becomes smaller as o decreases, until it vanishes at critical value $o_c(q) > 0 \; \forall q \geq 2$, whose exact value depends on q, as shown in Fig. 9.2(c). This implies that for low structural overlap globalisation is not achievable, and the system always has a qualitative different behavior, always converging towards a multicultural state independently of the number of cultural traits $q \geq 2$. This is in agreement with empirical finding of fragmented social states even in the presence of a limited number of available cultural choices.

A symmetric situation is observed if we study the system as a function of o, as shown in Fig. 9.2(d), where it is evident that when q is small fragmentation can be obtained only

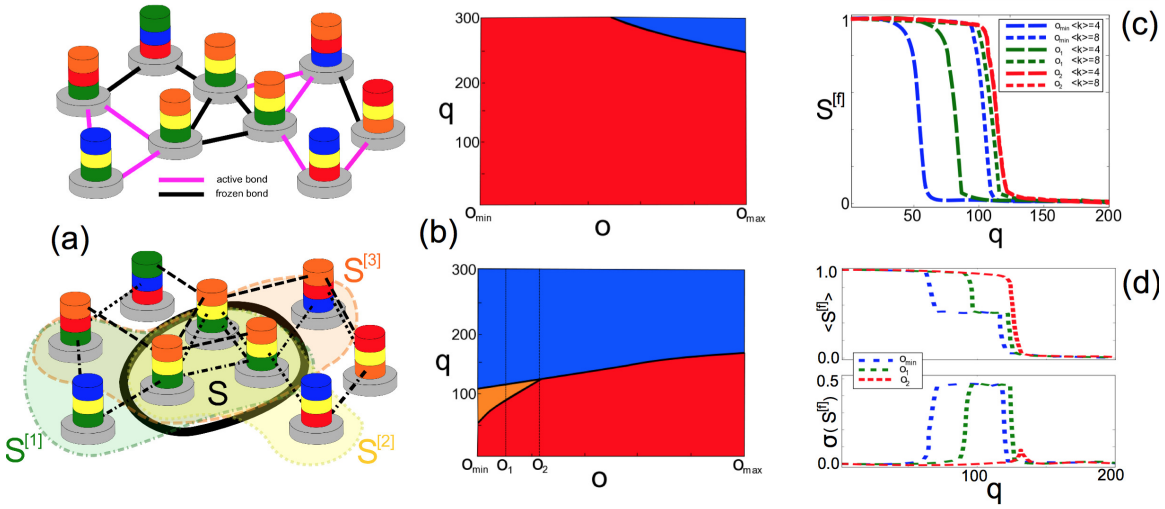


Figure 9.3: **Global and feature-level consensus.** A toy social system with three cultural features (a). With classical social influence (top), such system is still far from the equilibrium. Conversely, with layered social influence (bottom), the system is already in the absorbing state, and the largest cultural components at each layer differ from the global one. In (b) we show the phase diagram of the global largest cultural component of the system for classical (top) and layered (bottom) social influence as a function of q and for different values of the overlap o in networks with $N = 625$, $F = 10$ and five low-density layers ($\langle k \rangle = 4$) and five high-density layers ($\langle k \rangle = 8$). In the classical case, the system reaches a state where either globalization (red region) or fragmentation (blue region) is achieved for all features. In the layered case, instead, when o is smaller than a critical value o_2 , it is possible to have globalisation in the five high-density layers together with multiculturality in the five low-density ones (orange region). In (c), for the layered case, we plot the values of S as a function of q for one low-density and one high-density layer, and for three different values of the overlap, namely $o_{\min}$, o_1 and o_2. In (d), for the same values of overlap, we show the average and the standard deviation of the largest component $S^{[f]}$ over all layers [198].

for very low values of structural overlap. As q increases, the globalised phase shrinks and the critical value o_c increases. For $q > 140$, finally, the system is always in a fragmented state, even in the case of maximal structural overlap.

A important feature of social systems is the presence of cultural drift, i.e., the occasional spontaneous mutation of a cultural trait of an agent. Such phenomenon can be modelled as a noise of constant rate r, acting on the system on longer time-scales compared to the one that regulate the imitation and interactions among individuals. In the

classical Axelrod model on finite populations the multicultural absorbing states (blue region in Fig. 9.2(a)) are known to be metastable and fragmentation is destroyed if spontaneous mutation is allowed [208, 209], still leaving unanswered how it is possible to explain the persistence of multiculturality in realistic societies.

In Fig. 9.2(b) we show with different dashed lines the position of the critical line separating globalisation from fragmentation under four increasing values of noise, namely $r_1 = 3 \times 10^{-5}$, $r_2 = 5 \times 10^{-5}$, $r_3 = 6 \times 10^{-5}$ and $r_4 = 7 \times 10^{-5}$. We note that, given a number of cultural traits q, the critical value of overlap needed to sustain globalisation depends on the amount of noise in the system, i.e. $o_c = o_c(q, r)$. In analogy to the classical case, for $o > o_c(q, r)$ the multicultural region appearing at high values of q is unstable under cultural drift. Conversely, the polarised phase obtained for $o < o_c(q, r)$ is not affected by the presence of noise and allows to explain the persistence of multicultural states that we observe in real-world societies. Noticeably, in the limit $q \to 2$ the value of $o_c(q, r)$ appears to be independent of r and approaching the value of $2/F$. Conversely, in the classical model shown in 9.2(a) already all multicultural states for $q < 200$ are destroyed already for a drift as small as $r_2 = 5 \times 10^{-5}$, and all multicultural states for $q < 300$ are lost for $r_3 = 6 \times 10^{-5}$, no matter the value of the structural overlap. We remark that the instability of multicultural states is meant under the introduction of a moderate level of drift. Conversely, if drift is too high, the system enters a noisy dynamical state, describing an unrealistic society where spontaneous mutation accounts for most of the cultural chances and the effect of social influence is widely neglected.

9.4 Feature-level consensus

We introduce now the largest topical cultural component $S^{[f]}$. Two individuals i and j are considered to be part of the same topical component f if there is a path connecting them at layer f, and share the same cultural trait on feature f, i.e. $s_{f(i)} = s_{f(j)}$ [198]. We note that when $S = 1$, the condition that $S^{[f]} = 1 \ \forall f$ is automatically satisfied in both the classical and the layered model (assuming that each layer is connected). A remarkable property of layered social influence is that it allows for different levels of consensus on single topics to coexist, possibly resulting in the emergence of globalisation on some topics, i.e. $S^{[f]} \approx 1$ for some features f, and the persistence of fragmentation in some other ones, $S^{[\tilde{f}]} \approx 0$ for some other $\tilde{f}$. We call this regime *feature-level consensus*, or *topical consensus*. We remark that this realistic phenomenon can not be achieved with the classical version of the model, where at the absorbing state the largest component on each layer is always as large as the one computed taking into account all topics simultaneously.

In Fig. 9.3 we show how feature-level consensus may be possible only as an effect of the difference in the number of social interactions related to the different features, i.e. different layers' densities [198]. Let us consider two sets of layers. The first set is made of five equal layers with average degree $\langle k^{[f]} \rangle = 8$, whereas the second is made of five equal layers with average degree $\langle k^{[f]} \rangle = 4$. When all edges of the layers in the second set also belong to those in the first set, the structural overlap in the system is maximised, $o = o_{\max}$. Conversely, when no edges in the low-density layers are also present in the high-density ones the overall will be minimum, $o = o_{\min}$, and the system behaves as two isolated five-layer classical models with different average degree. Hence, because of the different density, they will have a different critical values q_c for transitioning from globalisation to multiculturality. Intermediate configurations when $o_{\min} < o < o_{\max}$ are also considered.

For $o = o_{\min}$, the critical values separating globalisations from multiculturality for the low-density and high-density layers are different. While both sets of layers are globalised for $q < 60$ and fragmented for $q > 120$, in the range $60 < q < 120$ globalisation is achieved only in the high-density layers, for which $S^{[f]} \approx 1$, and the low-density ones are instead polarised, $S^{[f]} \approx 0$. Strikingly, such property of the system is not peculiar of the case $o = o_{\min}$, but it is preserved for a finite range of values of structural overlap of to a critical value o_2. Such region is highlighted in orange in the diagram shown in Fig.9.3(b,bottom), and separates the globalised region (red) from the fragmented one (blue) when the value of overlap is sufficiently low.

The value of $S^{[f]}$ for low-density and a high-density layers is also shown in Fig. 9.3(c), as well as the average value of the largest cultural components $\langle S^{[f]} \rangle$ and its standard deviation. The standard deviation $\sigma(S^{[f]})$, approximately 0 for both the globalised and fragmented regions, takes values close to 0.5 in the mixed orange region, where consensus only exists on some layers.

9.5 Applications to real-world social networks

We remark that higher heterogeneity in the structure of the layers produces even richer and more diverse patterns of consensus across the different cultural features. A typical case is when the activity of the nodes on the layers is heterogenous [22], i.e. when not all agents are involved in discussion with other individuals on all topics, preventing the spread of a cultural traits across the whole population for some given features. This is the case of many real-world systems, where the structure of interactions is often inherently layered and diverse levels of activity occur at the different layers. As a test-

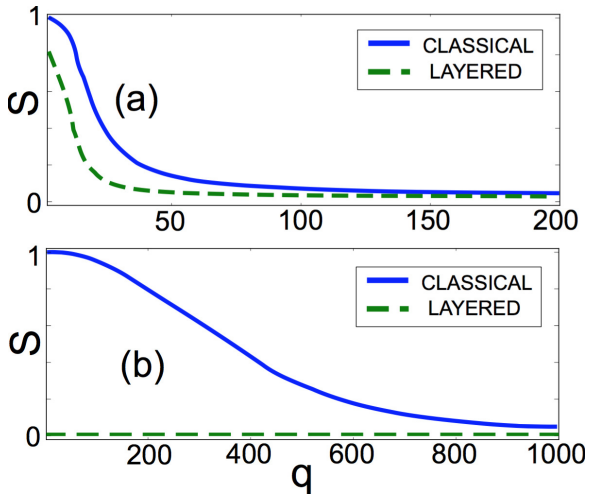


Figure 9.4: **Multiculturality in empirical multilayer social networks.** The largest cultural component S for classical (solid lines) and layered social influence (dashed lines) is reported as a function of q for two real-world multiplex social systems, respectively the network of Indonesian terrorists in panel (a) and the collaboration network of the Pierre Augier observatory in panel (b). In the first case the outcome of the classical and layered model are similar, with layered social influence only producing a decrease in the value of q_c. This is mainly due to the relatively large amount of structural overlap in the terrorists network ($o = 0.48$) and low heterogeneity in the activity of nodes. Conversely, the low value of redundancy in the Pierre Augier collaboration network ($o = 0.07$), and the heterogeneity in the density and activity of the layers, produce qualitatively different results: the layered model is always fragmented, even at $q = 2$, whereas the classical model produces a largest component of order $1/N$ only for $q > 1000$ [198].

case, we study the dynamics of our model with layered social influence on two multilayer social systems, namely the network of Indonesian terrorists introduced in Chapter 1 and extensively analysed in Ref. [15], and the collaboration network of the Pierre-Auger observatory presented in Ref. [126]. We recall that in the first dataset most of the 78 individuals are active on all $F = 3$ layers (as in most of this thesis, here we do not consider the fourth extremely sparse layers with information about financial ties), the multiplex network is connected and we have a structural overlap equal $o \approx 0.4817$. The second dataset consists of $F = 16$ layers and is in principle disconnected. In order to allow, at least potentially, complete globalisation also in this system, we selected the subset of individuals who belong to the giant component of the aggregated network. Such system includes 475 individuals. The system has a structural overlap of $o \approx 0.069$ and high heterogeneity in the activity of its members.

In Fig.9.4 we plot for both systems, as a function of q, the size of the largest cultural components found in the model with layered social influence (dashed lines) compared with those obtained with the classical model simulated on the corresponding aggregated networks (solid lines). For the terrorists network (panel (a)), a giant component of the order of the number of nodes appear for small q even in the layered case, while for $q \approx 20$ fragmentation appears on all features. The classical model predicts instead $q_c \approx 50$. For the Pierre-Auger collaboration network (panel (b)), in the classical model we have $S \approx 1/N$ only for $q > 1000$. Conversely, the layered model predicts multiculturality already for $q = 2$, producing a qualitative different result. Such striking difference between the layered and classical dynamics is due to the relatively low level of structural overlap and to the heterogeneity in the activity and density of the layers of the multiplex [198]. Such qualitative different behaviors highlights the importance of the underlying structure of social interactions to understand cultural patterns in real-world societies.

9.6 Discussion

Reconciling the phenomenon of social influence and imitation with the empirical evidence of multiculturality at a global scale is a long-debated problem. The model for the dissemination of culture introduced by Axelrod in 1997 [179] has proven to be able to produce socially diversified states in spite of the existence of a locally polarising rule. However, such multicultural states occur only for a very high number of cultural traits q and are not robust against cultural drift. In spite of several proposals [210, 218, 219, 224, 225], explaining the emergence and persistence of multicultural states that we experience in our life still appears as an open problem. In our work we considered agents having different interaction patterns according to the different cultural features. This is a property of many social systems, where individuals usually choose the recipients of their discussion according to the content of their messages and vice versa. By constraining the imitation of the cultural traits to the cultural features where two individuals are actually linked, we naturally introduce the notion of layered social influence, inherently connected to the average structural overlap o of a social system. Such definition implies the existence of stable states where connected agents might not have completely equal or different cultural profiles, but instead share just a limited number of features.

A main finding is that a robust multicultural regime emerges as a natural consequence of taking into account the actual structure of social interactions, which are inherently layered. Interestingly, when the structural overlap among the layers is small, i.e. when connections are not too much redundant across layers, only multicultural states will be allowed, independently of the actual number of possible cultural traits, suggesting a

qualitative shift in the behavior of the system. This finding is also in agreement with the empirical evidence of socially fragmented societies even under a limited number of cultural choices. Moreover, differently from the multicultural regime predicted when the layering of social influence is not taken into account, and achievable only for a large number of traits, such novel multicultural regime is robust to the presence of cultural drift, providing a potential explanation to the persistence of global diversity over time.

Another interesting aspect of layered social influence is the ability to explain the existence of a novel regime where globalisation appears only on a limited number of cultural features, while the remaining ones are characterised by multiculturality. The coexistence of different levels of multiculturality in different aspects of a cultural profile is in fact another typical property observed in human societies.

In conclusion, the model of layered social influence proposed in this paper, provides a mechanism leading to robust multiculturality, reveiling the important role played by a multilayer organisation of social interactions in avoiding globalisation, and pointing out the effect of heterogeneity of layer densities in the emergence of partially globalised regimes. While previous proposals to account for robust multiculturality invoke different forms of interactions among agents in an aggregated social network, we provide here an alternative mechanism not based in the form of the interaction, but on the structural properties of the social network. From this new perspective, the relevant control parameter of polarization-globalization transitions is no longer the number of cultural traits per feature, but a structural property, namely the overlap parameter accounting for the heterogeneity of social links among the different layers of the social network. The relevance of the structural properties of multilayared social interactions, illustrated here for cultural dissemination, is far reaching, and should be taken into account to reconsider a number of results based on models of social interactions among agents in aggregated networks.

Chapter 10

Evolutionary dynamics of group interactions

In the previous Chapter we considered a social dynamics where agents are influenced by their acquaintances through social influence. An individual selected a first neighbour at random and updated its state by imitating that of the other agent. Despite imitative processes are very frequent in nature, in many contexts imitation does not occur at random, but according to strategic behavior whose aim is to maximise some form of fitness or payoff. The field where game theory is applied to evolving populations is known as evolutionary game theory, whose synergies with statistical physics have significantly improved our understanding of public cooperation in structured populations. The interplay between the overlap in the structure of the layers and the control parameters of the corresponding games on sustaining cooperation in a multiplex networks has not yet been investigated. With this aim, we consider here a public goods game on a multiplex network, and we unveil the role of number of layers, and overlap of links and presence of uneven synergy factors at the different layers on the onset of cooperation [229]. In particular, we show that enhanced public cooperation emerges when a significant edge overlap is combined with at least one layer being able to sustain some cooperation by means of a sufficiently high synergy factor. In the absence of either of these conditions, the evolution of cooperation in multiplex networks is determined by the bounds of traditional network reciprocity with no enhanced resilience. These results caution against overly optimistic predictions that the presence of multiple social domains may in itself promote cooperation, and they help us better understand the complexity behind prosocial behavior in layered social systems.

10.1 Evolutionary game theory on structured populations

Human cooperation is an evergreen puzzle [230], at the heart of which is the divide between the Darwinian desire to maximize personal benefits and our social instincts that dictate prosocial behavior. The later are particularly strong in humans, because without their evolution we would have had serious challenges in rearing offspring that survived [231], and as a results would have likely died out as a species. Instead, we have acquired remarkable other-regarding abilities that have propelled us to dominance over all the other animals, to the point where today the biggest threat to us is ourselves.

The theoretical framework used most frequently to study cooperation among selfish individuals is evolutionary game theory [232–236], where the concept of a social dilemma captures the essence of the problem. In short, cooperation is costly, and it therefore weighs heavily on individual wellbeing and prosperity. One is thus torn between doing what is best for the society, and doing what is best for oneself. The public goods game is particularly apt in describing the dilemma [237]. The game is played in groups, where individuals can decide between cooperation and defection. Those that decide to cooperate contribute an amount to the common pool, while defectors contribute nothing. All the contributions are multiplied by a synergy factor that takes into account the added value of a group effort, and the resulting public goods are divided equally among all group members irrespective of their strategy. Clearly the best individual strategy is defection. But if everybody decides to defect there will be no public goods. In order to avoid the tragedy of the commons in a society cooperation is thus needed [238].

While the evolution of cooperation has been studied at great lengths in biology and sociology [239, 240], the problem became attractive for physicists after the discovery of network reciprocity [241], which manifests as the formation of resilient cooperative clusters in a structured population. Cooperators in the interior of such clusters can survive at conditions that do not sustain cooperation in well-mixed populations. In fact, methods of statistical physics have recently been applied to subjects that, in the traditional sense, and in particular evolutionary games in structured populations [242–244].

We are here concerned in particular with the evolution of cooperation in multilayer networks [146]. Several mechanisms have already been discovered by means of which the interdependence between different networks or network layers may help to increase the resilience of cooperation and resolve social dilemmas [245–251]. Interdependent network reciprocity is one example, which requires simultaneous formation of correlated cooperative clusters on two or more networks [249]. Other mechanisms that promote cooperation

beyond the bounds of traditional network reciprocity include non-trivial organization of cooperators across the network layers [246], probabilistic interconnectedness [248], information transmission between different networks [250], as well as self-organization towards optimally interdependent networks by means of coevolution [251].

Previous research has thus shown that multiplex networks may enhance the resilience of cooperation, but the key determinants of this, especially in terms of the topological overlap between the network layers and the game parametrization on each individual layer, still need to be determined. By studying the public goods game in a multiplex of regular random graphs, we here show that enhanced public cooperation requires significant edge overlap, combined with at least one layer being able to sustain some cooperation by means of a sufficiently high synergy factor [229]. The details of this conclusion depend further on the number of layers forming the multiplex, and on other properties of the spatiotemporal evolutionary dynamics, which includes pattern formation and spontaneous symmetry breaking across the layers. As we will show, these results provide a deeper understanding of the complexity behind cooperation in multiplex networks, and as such they have important implications for promoting prosocial behavior in different but linked social contexts.

10.2 The multiplex public goods game

In the public goods game players, belonging to a group of size G, are asked to contribute to a common pool. Cooperators contribute with a token d, typically $d = 1$, whereas defectors do not contribute at all. The amount of tokens in the pool is multiplied by a synergy factor r, and the resulting amount is divided equally among all players. Cooperators thus obtain a payoff $d(N_C \cdot r/G - 1)$, while defectors get $dN_C \cdot r/G$, where N_C is the number of cooperators in the group. When the game is played on multiple rounds, players choose to cooperate or defect at each iteration, based on the success of the two strategies. The game can be implemented on structured populations, where players are placed on the nodes of a graph and interact through their links. The results usually show that, while in well-mixed populations cooperators extinguish quickly, repeated local interactions among the same players allow the formation of clusters of cooperators which are able to survive.

In real situations, individuals are typically involved in strategical choices on independent domains, and can adopt different strategies according to the specific domain. However, information on the earnings of each individual might be only be available at the aggregate level, as a sum the payoff obtained as a result of all its decisions. More

formally, in order to take this into account, we consider here a population of N individuals playing the public goods game on the M layers of a multiplex network. In particular, we model each layer as a regular random graph with degree $k = 4$. Hence the game is played in groups each of size $G = k + 1$. The state of a player i, $i = 1, 2, \ldots, N$, is fully described by a vector of strategies $\boldsymbol{\sigma}_i = \{\sigma_i^{[1]}, \ldots, \sigma_i^{[M]}\}$, such that at each layer α, $\alpha = 1, 2, \ldots, M$, the player can independently choose to either cooperate, i.e., $\sigma_i^{[\alpha]} = +1$, or defect, i.e., $\sigma_i^{[\alpha]} = -1$. The benefit of synergy among individuals in general depends on the specific domain. In order to model this feature, we assume that the synergy factor can be different from layer to layer. We hence consider a synergy factor vector $\boldsymbol{r} = \{r^{[1]}, \ldots, r^{[M]}\}$. On each layer α, player i earns the payoff $\pi_i^{[\alpha]}$, such that i gains $d(N_C \cdot r^{[\alpha]}/G - 1)$ if it cooperates, or otherwise $dN_C \cdot r^{[\alpha]}/G$, if it defects.

The public goods game is simulated by a Monte Carlo method in which, at each elementary step, a layer α is selected, and then a randomly chosen node i, and one of its neighbours j on that layer, are considered. Both i and j play the game on all the layers, thereby obtaining respectively payoffs $\pi_i = \sum_{\alpha=1}^{M} \pi_i^{[\alpha]}$ and $\pi_j = \sum_{\alpha=1}^{M} \pi_j^{[\alpha]}$. Finally, player i compares its payoff to that of player j, and copies the strategy of player i, but only at the layer α, with a probability given by a Fermi function:

$$W(\sigma_i^{[\alpha]} \to \sigma_j^{[\alpha]}) = \left(1 + exp\left[\frac{\pi_j - \pi_i}{K}\right]\right)^{-1}, \tag{10.1}$$

where K quantifies the contribution of random fluctuations to the strategy adoption [252–254]. In the $K \to 0$ limit, player j copies the strategy of player i if and only if $\pi_i > \pi_j$. Conversely, in the $K \to \infty$ limit, payoff differences cease to matter and i copies the strategy of j with a probability equal to 0.5. Between these two extreme cases, for intermediate values of K, players with a higher payoff will be readily imitated, although the strategy of under-performing players may also be occasionally adopted to mimic, for example, errors in the decision making, imperfect information and external influences that may adversely affect the evaluation of an opponent. We adopt the value $K = 0.5$ without loss of generality, as shown in [253]. In our simulations, we obtain one full Monte Carlo step (MCS) by repeating $M \times N$ times the elementary steps described above, thus giving a chance to every player to change its strategy on all the layers once on average.

In order to characterize the outcomes of our evolutionary dynamics model, we introduce the vector $\boldsymbol{c} = \{c^{[1]}, \ldots, c^{[M]}\}$, where $c^{[\alpha]}$ is the fraction of cooperators at layer α in the stationary state, i.e., when the average over time of this quantity becomes time

independent. As a first order parameter we consider then the quantity

$$c = \frac{1}{M}\sum_{\alpha=1}^{M} c^{[\alpha]}, \tag{10.2}$$

which is the overall fraction of cooperators across all the layers of the multiplex in the stationary state. We note that $0 \leq c \leq 1$, where $c = 1$ corresponds to full cooperation while $c = 0$ corresponds to full defection. As a second order parameter, we define the average coherence ξ of the players across all the layers, defined as:

$$\xi = \frac{1}{N}\sum_{i=1}^{N} \xi_i \qquad \xi_i = \left| \frac{1}{M}\sum_{\alpha=1}^{M} \sigma_i^{[\alpha]} \right| \tag{10.3}$$

where ξ_i is the coherence of the strategies of player i. A value $\xi_i = 1$ indicates that player i is maximally coherent, meaning it adopts the same strategy in all the layers. Conversely, $\xi_i = 0$ means that player i is maximally incoherent, adopting $\sigma_i^{[\alpha]} = +1$ just as often as $\sigma_i^{[\alpha]} = -1$ across the M different layers. A similar definition of coherence for the particular case $M = 2$ has been reported in Sec. 8.2. In addition to the two main order parameters c and ξ, when $M = 2$, we will also use the quantities $c^{[2]} - c^{[1]}$ and $|c^{[2]} - c^{[1]}|$ to evaluate differences in the level of cooperation at the two layers. In fact, as we will show in the following, there exist indeed regions in the parameter space of our model such that full cooperation is observed at the layer with the highest synergy factor, while the layer with the lower synergy factor is in a full defection state. Also, a multiplex network with two layers may exhibit spontaneous symmetry breaking, such that, even if the two layers are characterized by the same synergy factor and the same interaction network topology, the level of cooperation on them can be different.

10.3 Structural effects: number of layers and edge overlap

We implement our model on a multiplex network in which it is possible to tune the similarity among the topology of the M layers. We therefore consider that each layer is a regular random graph with $N = 10^4$ nodes and $K = 2 \cdot 10^4$ links, and we tune the average edge overlap ω defined as in Eq. 2.10,where the average is restricted to the pairs of nodes which share at least one edge. We begin by assuming that the synergy factor used for the public goods game is the same at each layer, namely we set $r^{[\alpha]} = r \quad \forall \alpha$. Initially, each layer is populated by the same proportion of cooperators and defectors, distributed uniformly at random, and subsequently the game is iterated in time according to the Monte Carlo simulation procedure described in Section II. Results presented in

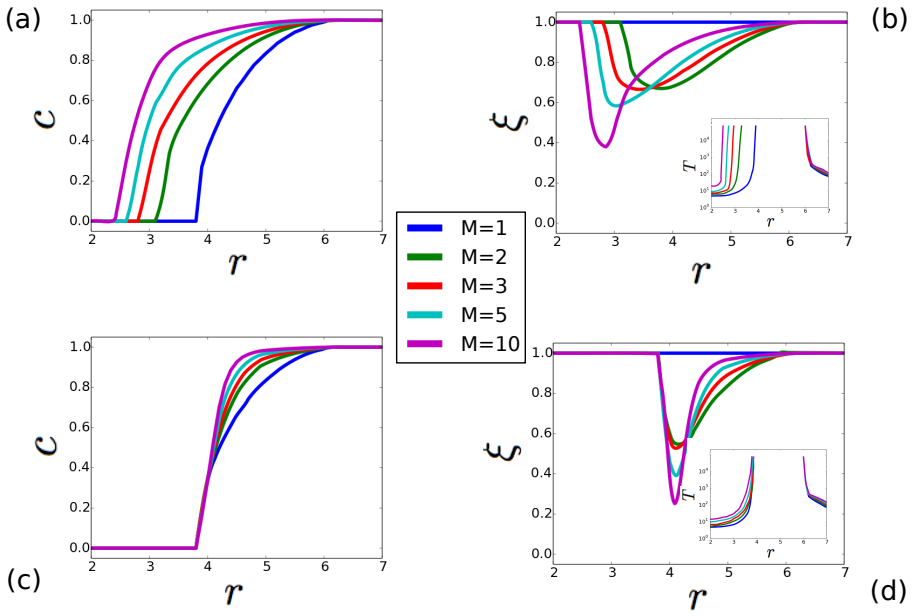


Figure 10.1: Number of layers and topological overlap are crucial to lower the critical value of the synergy factor needed for cooperators to survive. (a,b) The multiplex is formed by M regular random graphs with degree $k = 4$, and edge overlap respectively equal to $\omega = 1$ (a,b) and $\omega = 0$ (c,d). We show the average fraction of cooperators across the whole multiplex c (a,c) and the average coherence of the players across all the layers ξ (b,d) as a function of the synergy factor r, and for different values of M. Insets show the number of full Monte Carlo steps T needed for the system to reach an absorbing phase with either all cooperators or all defectors [229].

Fig. 10.1 are shown separately in two rows, respectively for the case $\omega = 1$ [panels (a) and (b)] and the case $\omega = 0$ [panels (c) and (d)]. Looking at panels (a) and (b), it can be observed that the larger the value of M, the lower the critical value of the synergy factor that is needed to sustain cooperation. When $M = 1$, on a single-layer regular random graph, the critical value is equal to $r_c = 3.75$, which is in agreement with traditional network reciprocity [253]. When ten layers form the multiplex, however, the critical value drops to as low as $r_c = 2.35$. The minimal coherence also emerges at ever lower values of r as M increases, and the minima become lower, indicating that at least some layers are able to sustain cooperation even though in the majority the players defect [229].

The evolutionary outcomes are significantly different in panels (c) and (d), where the topological overlap is zero. It can be observed that the increase in M does nothing to reduce the critical values of r needed to sustain cooperation. In fact, $r_c = 3.75$ that is due to traditional single-layer network reciprocity always emerges as the necessary condition

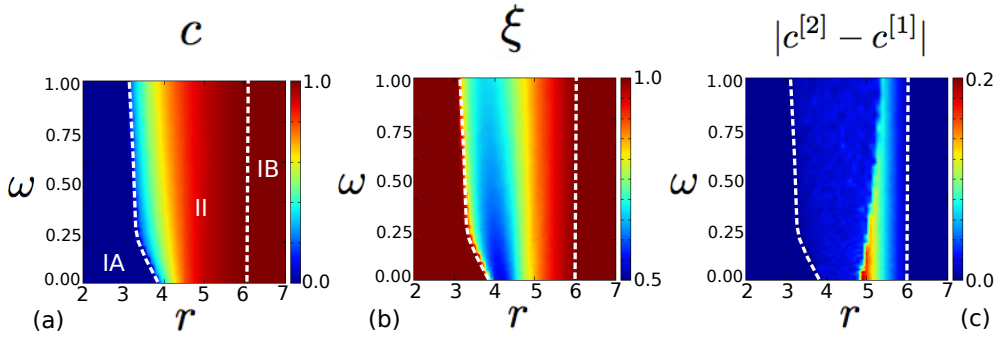


Figure 10.2: Benefits to public cooperation in a multiplex network with $M = 2$ layers and a tunable value of edge overlap ω. The synergy factor used for the public goods game is the same at the two layers, namely $r^{[1]} = r^{[2]} = r$. Panel (a) shows the full $r - \omega$ phase diagram where the color map encodes the average fraction of cooperators c across the two layers. High benefits emerge only for large values of ω. Panels (b) and (c) show the full $r - \omega$ phase diagram where the color map encodes respectively the average coherence ξ and the absolute difference $|c^{[2]} - c^{[1]}|$ between the fraction of cooperators in the two layers. Reported dashed lines separate regions where the system is at an absorbing state with full coherence $\xi = 1$ and either full defection $c = 0$ (IA) or full cooperation $c = 1$ (IB), from the region of continuously evolving coexistence of cooperators and defectors $0 < c < 1$ (II) [229].

for public cooperation. Not surprisingly, the minimal coherence also occurs at the same value or r regardless of M. The minima become lower as M increases because the average goes over more layers, simply giving statistically more opportunity for players to hold different strategies across different layers. Taken together, these results show that topological overlap is essential for enhanced multiplex network reciprocity to take effect and enhance the resilience of public cooperation expected in a system with multiple layers of interactions [229]. These results confirm the crucial impact of the edge overlap on dynamical processes on networks, as we already saw in the previous Chapter for the Axelrod model.

To investigate more in details the role of the topological overlap, we consider a multiplex with only two layers, but where the value of the edge overlap ω can be varied continuously in the range $[0, 1]$ through the model introduced in Sec. 4.2. In Figure 10.2 we report the results obtained as a function of the two control parameters r and ω. The three phase diagrams shown encode respectively the average fraction of cooperators c across the two-layer multiplex (a), the average coherence ξ of the players (b), and the absolute difference between the fraction of cooperators in the two layers, $|c^{[2]} - c^{[1]}|$ (c). In panels (a) and (b), we can distinguish two regions, namely type I where the

whole multiplex reaches an absorbing phase (each layer is either in full cooperation or in full defection), and type II where the multiplex is trapped in a state where cooperators and defectors coexist. The two type I regions can be further classified as type IA where defectors dominate ($c = c^{[1]} = c^{[2]} = 0$), and type IB where cooperators dominate ($c = c^{[1]} = c^{[2]} = 1$). In both type IA and IB regions all the players are of course fully coherent, i.e., they adopt the same strategy on both layers such that $\xi = 1$. It can be observed that the added value of the multiplex structure in enhancing network reciprocity marking the transitions from region IA to II suddenly decreases for $\omega < 0.25$, disappearing as the value of the topological overlap ω approaches 0 [229].

Interestingly, at the II to IB transition, that is from the mixed $(C + D)$ phase to the pure C phase, the topological overlap does not play a role at all, indicating that the enhanced multiplex network reciprocity is crucial only when cooperation can be barely sustained. Even as the multiplex enters the mixed $(C + D)$ phase, i.e., region II, the impact of the extent of topological overlap vanishes very quickly beyond the critical value of r_c at the transition point. In the mixed $(C + D)$ phase, we can also observe spontaneous symmetry breaking in panel (c), where in region II $|c^{[2]} - c^{[1]}| > 0$. This means that, even though the public goods game in both layers is characterized by the same synergy factor and is staged on layers with identical topological properties, the level of cooperation in the stationary state is different. In particular, it can be observed that the lower the topological overlap between the two layers ($\omega \to 0$), the higher the symmetry breaking, with the maximum value occurring for $\omega = 0$ and $r = G = 5$ [229].

10.4 Dynamical effects: the role of different synergy factors

Lastly, we study the impact of differing synergy factors in the layers forming the multiplex in order to determine the importance of game parametrization on the emergence of enhanced multiplex network reciprocity. In Fig. 10.3, we present results separately for two-layer multiplex networks with complete (a,b,c) and zero (d,e,f) topological overlap between the layers. The $r^{[1]} - r^{[2]}$ phase diagrams encode the average fraction of cooperators across the two-layer multiplex c (a), the average coherence of the players across the two layers ξ (b), and the difference between the fraction of cooperators in the two layers $c^{[2]} - c^{[1]}$ (c). It can be observed that in both cases, regardless of the overlap, region IB occurs when both $r^{[1]} > 6$ and $r^{[1]} > 6$ (a,d). A new region can also be observed in both cases when $r^{[\alpha]} > 6$ and $r^{[\beta]} < 3.75$, which we denote as IC, where one layer is characterized by full cooperation ($c^{[\alpha]} = 1$), while the other layer is characterized by full

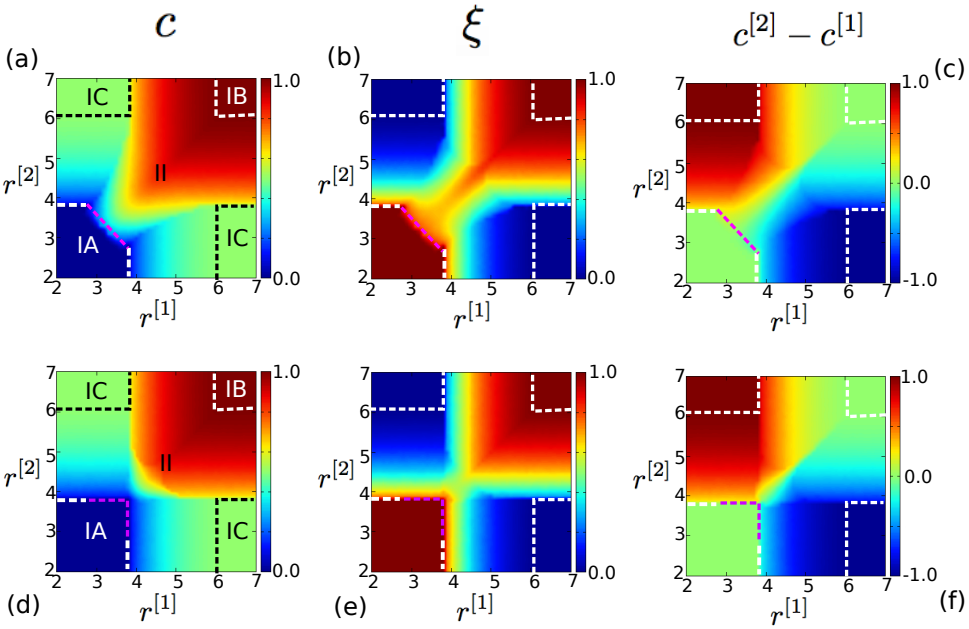


Figure 10.3: The emergence of multiplex network reciprocity depends on the different values of the synergy factors at each layer. The multiplex is formed by two layers of regular random graphs with an edge overlap respectively equal to $\omega = 1$ (a,b,c) and $\omega = 0$ (d,e,f). Panels (a,d) report the full $r^{[1]} - r^{[2]}$ phase diagram where $r^{[1]}$ and $r^{[2]}$ are the synergy factors at the two layers, and where the color map encodes the average fraction of cooperators c. Panels (b,e) and (c,f) show the full $r^{[1]} - r^{[2]}$ phase diagram where the color map encodes respectively the average coherence ξ and the difference of cooperators $c^{[2]} - c^{[1]}$ in the two layers. A new absorbing state (region IC), with full cooperation at layer α $c^{[\alpha]} = 1$, full defection at layer β $c^{[\beta]} = 0$ and complete incoherence $\xi = 0$, emerges [229].

defection ($c^{[\beta]} = 0$). Accordingly, we have a completely incoherent multiplex with $\xi = 0$ (b,e).

These equivalences beget the question when do the evolutionary outcomes actually differ in dependence on complete and zero overlap. As results in Fig. 10.3 show, and as could be anticipated from the results presented in Fig. 10.2, the difference is most expressed at the interface between regions IA and II (see dashed purple line). If there is no topological overlap between the two layers (d,e,f), we see that as long as both $r^{[1]} < 3.75$ and $r^{[2]} < 3.75$, we always have full defection in both layers. Hence, multiplexity does not provide any advantage to the evolution of cooperation (see also Fig. 10.1). Conversely, when there is perfect overlap (a,b,c), cooperators emerges already for $r = r^{[1]} = r^{[2]} < 3.75$, roughly $r \approx 3.25 = r_c$. But given an arbitrary choice for $r^{[\beta]}$ that is smaller

than 3.75, is $r^{[\alpha]} = 3.25$ the minimum value to see the emergence of cooperators in the multiplex? Indeed no, given $r^{[\beta]}$ we still see cooperators in the system as long as we choose $r^{[\alpha]}$ such that it is slightly above $\frac{r^{[\alpha]}+r^{[\beta]}}{2} > 3.25$ (linear relationship $r_c^{[\alpha]} = 2 \cdot 3.25 - r^{[\beta]}$, see purple line). Importantly, this relation holds as long as $r^{[\beta]}$ does not go beyond 3.75, the original critical value for one layer, at which point the relation no longer holds and the overlapping case behaves as the non-overlapping case. Based on the phase diagrams in Fig. 10.3, the critical value can be approximated as $r^{[\alpha]} \approx 2 \cdot r_c - r_c(M = 1)$, which means $r^{[\alpha]} = 2 \cdot 3.25 - 3.75 \approx 2.75$. Taken together, a topologically overlapping multiplex can extend the coexistence region II towards significantly smaller values of r, which in our two-layer setup corresponds to a triangle delimited by $(r^{[1]}, r^{[2]})$ such that $X = (2.75, 3.75)$, $Y = (3.75, 3.75)$ and $Z = (3.75, 2.75)$ [229].

10.5 Discussion

We have studied the determinants of public cooperation in multiplex networks, focusing in particular on the topological overlap and different synergy factors across the layers. We have shown that, if the topological overlap between the layers is sufficiently extensive, the critical value of the synergy factor that enable cooperators to survive decreases steadily as the number of layers increases. This result confirms the existence of interdependent or multiplex network reciprocity, which enhance the resilience of cooperators beyond the bounds of traditional network reciprocity on an single-layer network. However, we have also shown that, as the topological overlap between the layers decreases, so do the benefits of multiplexity for the evolution of cooperation. In particular, if the topological overlap is zero, cooperators loose all benefits stemming from their engagement in different layers of the multiplex and thus become reliant on single-layer network reciprocity alone. These results manifest not only in the average fraction of cooperators in the multiplex, but also in the average coherence of the players across all the layers. We show that, in case of perfect topological overlap, the later reaches a minimum at ever lower values of the synergy factor as the number of layers increases, while in the absence of topological overlap the synergy factor yielding minimal coherence is independent of the number of layers.

By further varying the synergy factor that applies on each particular layer, we have shown that the topological overlap is crucial only if the synergy factor on all layers is smaller than the critical value on a single layer. If that is the case, the overlap plays a key role in sustaining cooperation, and there exists an average value of the synergy factor across all the layers that needs to be reached for cooperators to survive. However, if on a single layer the synergy factor is large enough to sustain cooperation even in the

absence of multiplexity, i.e., as if the layer would be isolated, then the topological overlap seizes to matter. By means of extensive Monte Carlo simulations, we have determined precise bounds on the topological overlap and the relations between synergy factors in different layers that need to be met for enhanced multiplex network reciprocity to take effect. Taken together, our results thus establish key determinants of public cooperation in multiplex networks.

The presented results reveal rather stringent conditions that have to be met for public cooperation to be more resilient on multiplex network than it is on single-layer networks. Indeed, the hallmark of multiplex network reciprocity, which is the simultaneous formation of correlated cooperator clusters across different layers, requires near perfect topological overlap, and is effective only if the conditions for cooperation on all layers are rather dire. If the coordination process leading to the formation of correlated cooperator clusters is disturbed due to the lack of topological overlap, multiplex network reciprocity never emerges, resulting in the total collapse of cooperation across all layers minus those that would sustain cooperation on their own either way. Thus, while multiplexity and network interdependence can in theory be exploited effectively to promote cooperation past the limits imposed by isolated networks, caution is needed against overly optimistic predictions that suggest involvement in different social contexts alone is in itself sufficient to promote cooperation. Enhanced prosocial behavior in layered social systems can emerge only if the positions of the players and the links among them in these layers do not differ much.

Conclusions

Network science is an old science: it draws on theories and methods from a variety of disciplines, such as graph theory from mathematics, statistical mechanics from physics, data mining from computer science, inference from statistics, and social network analysis from sociology, all of which have a very well developed trajectory of their own. Yet, network science is a new science: making all these fields speaking together it created the ground for a novel fertile research area, which has produced many interesting works in the last twenty years. The main reason behind its success is that its theoretical, numerical and mathematical tools were shown to provide, or contribute to find practical solutions to hard problems. For instance, the study of networks has helped us unveil the structure of the internet, of human and animal societies, and understand how epidemics and rumours spread or how epileptic seizures happen. Many complex systems are indeed well described in terms of networks, and their complexity typically dwells in the intricate structure of interactions among their units. In a quest to better represent and understand real-world systems, network scientists have progressively included realistic features in their network models and analyses, such as link weight, direction, sign and time. When links can be grouped in different types, we are dealing with multilayer, or multiplex, networks. Multiplex networks are everywhere, from multimodal transportation systems to multifaceted social relationships and biological interactions taking place through different channels. In the first part of this thesis we suggested measures and models to unveil and quantify multiplexity from a structural point of view, from basic node, layer and edge properties to more complicated structures at the micro- and meso-scale, such as motifs, communities and core-periphery. These tools are potentially of interest for many researchers working with networks, from sociologists to neuroscientists, economists, ecologists and urban scientists, and we collected most of them in a freely available and open source software library [255]. The second part of this thesis is devoted to dynamical processes on multiplex networks, and in particular random walks, opinion dynamics, cultural dynamics and evolutionary game theory. The crucial point is that interactions across layers may give rise to novel and unexpected emergent behavior, which is unpredictable from the knowledge of the single layers independently, or from

that of the aggregated system [33]. When this novel multiplex behavior exists, "multiplex matters" and the multiplex ingredient must not be missed in the representation of our complex system. The study of multiplex networks has received a lot of attention in the last years, almost as if they were a research topic of their own. Yet, no figure like that of a multiplex scientist must exist: the novel tools and techniques should just enter into the background and skills of researchers working on and with networks. Methods and results from multiplex network theory are likely to be combined in the upcoming years to other topics at the forefront of network science, such as temporal networks [8, 256] and networks with metadata [257, 258], combining multiple additional level of complexity. Together with the theory, we also hope and believe the study of multilayer systems will help bring answers to practical problems. Will multiplex network theory contribute to battle diseases by looking at their correlation and cooccurrence with specific habits? Will it be able to reduce the risk of dramatic interdependent cascades in the banking and in the financial sectors? Will multiplexity give us a better understanding of societies and of our brain? Hopefully this thesis represents a little step in generate research efforts in this direction, providing tools, methods and in general a new framework to analyse increasingly reacher data now available in fields as diverse as medicine and finance. For the moment, we can tell network scientists who are trying to unveil the structure of our brain, the world trade web, patterns of online and offline communication, and also all those kids who can't help stare at the Tube map in London while waiting for a train: yes, those are multiplex networks, and we understand your fascination.

References

[1] S. H. Strogatz, "Exploring complex networks", *Nature* **410,** 268–276 (2001).

[2] R. Albert and A.-L. Barabasi, "Statistical mechanics of complex networks", *Review of Modern Physics* **74,**, 47 (2002).

[3] M. E. J. Newman, "The structure and function of complex networks",*SIAM Review* **45,** 167-256 (2003).

[4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, "Complex networks: Structure and dynamics", D.U. Hwang, *Physics Reports* **424,** 175 (2006).

[5] D. Watts, S. Strogatz, "Collective Dynamics of Small-World Networks", *Nature* **393,** 440-442 (1998).

[6] A.L. Barabasi, R. Albert, "Emergence of scaling in random networks", **Science** **286** (5439), 509-512 (1999)

[7] A. Barrat, M. Barthelemy, R. Pastor-Satorras, A. Vespignani, "The architecture of complex weighted networks", *Proceedings of the National Academy of Sciences* **101**, 3747 (2004).

[8] P. Holme, J. Saramaki, "Temporal Networks", *Physics Reports* **519,** 3 (2011).

[9] S. Boccaletti, G. Bianconi, R. Criado, C. Del Genio, J. Gomez-Gardenes, M. Romance, I. Sendina-Nadal, Z. Wang, M. Zanin, "The structure and dynamics of multilayer networks", *Physics Reports* **544,** 1 (2014).

[10] M. Kivela, A. Arenas, M. Barthelemy, J. Gleeson, Y. Moreno, M. Porter, "Multilayer networks", *Journal of Complex Networks* **2,** 3 (2014).

[11] F. Battiston, V. Nicosia, V. Latora, "The new challenges of multiplex networks: measures and models" *The European Physical Journal Special Topics* **226** 3 401-416 (2017).

[12] M. De Domenico et al., "Mathematical formulation of multilayer networks", *Physical Review X* **3**, 041022 (2013).

[13] F. Radicchi, A. Arenas, "Abrupt transition in the structural formation of interconnected networks" *Nature Physics* **9** 11 717-720 (2013).

[14] G. Bianconi, "Statistical mechanics of multiplex networks: Entropy and overlap", *Physical Review E*, **87** 6 062806 (2013)

[15] F. Battiston, V. Nicosia and V. Latora, "Structural measures for multiplex net-

works" *Physical Review E* **89** 3 (2014).
[16] A. Cardillo et al., "Emergence of network features from multiplexity", *Scientific Reports* **3**, 1344 (2013).
[17] M. Szell, R. Lambiotte, S. Thurner, "Multirelational organization of large-scale social networks in an online world", *Proc Natl Acad Sci USA* **107** 31 13636-13641 (2010).
[18] M. Barigozzi, G. Fagiolo, D. Garlaschelli, "Multinetwork of international trade: A commodity-specific analysis", *Physical Review E* **81** 4 046104 (2010).
[19] L. Cantini, E. Medico, S. Fortunato, M. Caselle, "Detection of gene communities in multiplex networks reveals cancer drivers", *arXiv preprint arXiv:1507.08415* (2015).
[20] S. Pilosof, M.A. Porter, S. Kefi, "Ecological multilayer networks: A new frontier for network ecology", *arXiv preprint arXiv:1511.04453* (2015).
[21] R. Gallotti, M. Barthelemy, "Anatomy and efficiency of urban multimodal mobility", *Scientific Reports* **4**, 6911 (2014).
[22] V. Nicosia, V. Latora, "Measuring and modeling correlations in multiplex networks", *Physical Review E* **92** 3 (2015).
[23] M. De Domenico, V. Nicosia, A. Arenas, "Structural reducibility of multilayer networks", V. Latora, *Nature Communications* **6** (2015)
[24] S.V. Buldyrev, R. Parshani, G. Paul, H.E. Stanley, S. Havlin, "Catastrophic cascade of failures in interdependent networks" *Nature* **464**, 1025 (2010)
[25] S. Gomez et al, "Diffusion dynamics on multiplex networks", *Physical Review Letters* **110** 2 028701 (2013).
[26] N. Kouvaris, S. Hata, A. Diaz-Guilera, "Pattern formation in multiplex networks", *Scientific Reports* **5** 10840 (2015).
[27] C.I. Del Genio, J. Gomez-Gardenes, I. Bonamassa, S. Boccaletti, "Synchronization in networks with multiple interaction layers", *Science Advances* **2** (11), e1601679 (2016)
[28] J Sanz, C.Y. Xia, S. Meloni, Y. Moreno, "Dynamics of interacting diseases", *Physical Review X* **4** 4 041005 (2014).
[29] V Nicosia, P. Skardal, A. Arenas, V. Latora, "Collective phenomena emerging from the interactions between dynamical processes in multiplex networks", *arXiv preprint* arXiv:1405.5855 (2017)
[30] C. Granell, S. Gomez, A. Arenas, "Dynamical interplay between awareness and epidemic spreading in multiplex networks", *Physical Review Letters* **111** 12 128701 (2013).
[31] H. Wu, S. Gomez, A. Arenas, "On the influence of trust in the spreading of information", *arXiv preprint arXiv:11606.01688* (2016).
[32] R. Amato, A. Diaz-Guilera, K.-K. Kleineberg, "Interplay between social influence and competitive strategical games in multiplex networks", *arXiv preprint*, arXiv:1702.05952 (2017)
[33] F. Battiston, M. Diakonova, J. Gomez-Gardenes, V. Nicosia, V. Latora, "The emer-

gent dynamics of multiplex networks", in preparation, (2017)
[34] G. Menichetti, D. Remondini, P. Panzarasa, R. J. Mondragón, G. Bianconi, "Weighted multiplex networks", *Plos One* **9** (6), e97857 (2014).
[35] F. P. Ramsey, "On a Problem of Formal Logic", Proceedings London Mathematical Society, s2-30 (1) 264–286 (1930)
[36] P. Holme, "Modern temporal network theory: a colloquium", *The European Physical Journal B* **88**, (9), 1-30 (2015).
[37] P. Holme and J. Saramki, eds., "Temporal Networks" (Springer, Berlin), (2013)
[38] A. Saumell-Mendiola, M. Á. Serrano and M. Boguñá, "Epidemic spreading on interconnected networks" Physical Review E **86**, 026106 (2012).
[39] http://openflight.org
[40] http://journals.aps.org/datasets
[41] ftp://ftp.fu-berlin.de/pub/misc/movies/database/
[42] V. Colizza, A. Flammini, M.A. Serrano, A. Vespignani, "Detecting rich–club ordering in complex networks", *Nature Physics* **2**, 110 (2006).
[43] M. Newman, "The structure of scientific collaboration networks" *Proceedings of the National Academy of Sciences* **98** (2) 404–409 (2001)
[44] N. Roberts, S. F. Everton, *Roberts and Everton Terrorist Data: Noordin Top Terrorist Network (Subset)* (2011).
[45] R. Guimera, LAN Amaral, "Cartography of complex networks: modules and universal roles" *Journal of Statistical Mechanics: Theory and Experiment*, P02001 (2005).
[46] R. Guimera, LAN Amaral, "Functional cartography of complex metabolic networks", *Nature* **433,** 895-900 (2005).
[47] V. Latora and M. Marchiori, "Economic small-world behavior in weighted networks", *The European Physical Journal B* **32,** 249–263 (2003).
[48] S. Wasserman and K. Faust, *Social Networks Analysis*, (Cambridge University Press, Cambridge, 1994).
[49] E. Cozzo et al., "Structure of triadic relations in multiplex networks", *New Journal of Physics* **17** (7), 073029 (2015).
[50] M. G. Morris, M. Barthelemy, "Transport on coupled spatial networks", *Physical Review Letters* **109** 12 128703 (2012)
[51] A. Aleta, S. Meloni, Y. Moreno, "A multilayer perspective for the analysis of urban transportation systems" *arXiv preprint*, arXiv: 1607.00072 (2016)
[52] L. Solá, M. Romance, R. Criado, J. Flores, A. Garcia del Amo, S. Boccaletti, "Eigenvector centrality of nodes in multiplex networks", *Chaos: An Interdisciplinary Journal of Nonlinear Science* **23** (3), 033131 (2013)
[53] A. Solé-Ribalta, M. De Domenico, S. Gómez and A. Arenas, "Random walk centrality in interconnected multilayer networks", *Physica D*, **323**, 73-79 (2016)
[54] A. Halu, R. J. Mondragón, P. Panzarasa, G. Bianconi, "Multiplex pagerank", *PLoS One*, **8** (10) (2013)
[55] M. De Domenico, A. Solé-Ribalta, E. Omodei, S. Gómez, A. Arenas, "Ranking in

interconnected multilayer networks reveals versatile nodes", *Nature Communications* **6** (2015)

[56] A. Solé-Ribalta, M. De Domenico, S. Gómez, A. Arenas, "Centrality rankings in multiplex networks", *Proceedings of the 2014 ACM conference on Web Science* 149-155 (2014)

[57] K.-K. Kleineberg, M. Boguña, M. Angeles Serrano, F. Papadopoulos, "Hidden geometric correlations in real multiplex networks", *Nature Physics*, **12** 1076–1081 (2016)

[58] L. Lacasa, V. Nicosia, V. Latora, "Network structure of multivariate time series", *Scientific reports* **5** (2015)

[59] G. Ferraz de Arruda, E. Cozzo, Y. Moreno, F. A. Rodrigues, "On degree?degree correlations in multilayer networks", *Physica D*, **323**, 5-11 (2016)

[60] A. Solé-Ribalta, M. De Domenico, N. E. Kouvaris, A. Diaz-Guilera, S. Gómez, A. Arenas, "Spectral properties of the Laplacian of multiplex networks", *Physical Review E*, **88** (3), 032807 (2013)

[61] R.J. Sánchez-García, E. Cozzo, Y. Moreno, "Dimensionality reduction and spectral properties of multilayer networks", *Physical Review E*, **89** (5), 052815 (2014)

[62] R. Pastor-Satorras, A. Vazquez, A. Vespignani, "Epidemic spreading in scale-free networks", *Physical Review Letters* **87**, 258701 (2001)

[63] F. Battiston, J. Iacovacci, V. Nicosia, G. Bianconi, V. Latora, "Emergence of multiplex communities in collaboration networks", *PLoS One* **11** (1): e0147451. (2016)

[64] V. Gemmetto, D. Garlaschelli, "Multiplexity versus correlation: the role of local constraints in real multiplexes", *Scientific Reports*, **5** 9120 (2015)

[65] B.F. Skinner "The Behavior of Organisms: An Experimental Analysis", New York: Appleton-Century-Crofts (1938)

[66] C. J. Stam, et al., "Modern network science of neurological disorders", *Nature Review Neuroscience*, **15**, 683 (2014).

[67] A. Fornito, et al., "The connectomics of brain disorders", *Nature Review Neuroscience* **16**, 159 (2015).

[68] F. Varela, et al., "The brainweb: Phase synchronization and large-scale integration", *Nature Review Neuroscience* **2**, 229 (2001).

[69] G. Deco, et al., "Emerging concepts for the dynamical organization of resting-state activity in the brain", *Nature Review Neuroscience* **12** 43 (2011).

[70] V. Nicosia, et al., "Remote synchronization reveals network symmetries and functional modules", *Physical Review Letters* **110** 174102 (2014).

[71] C. J. Honey, et al., "Network structure of cerebral cortex shapes functional connectivity on multiple time scales", *Proceeding of the National Academy of Sciences* **104** 10240 (2007) .

[72] C. J. Honey, et al., "Predicting human resting-state functional connectivity from structural connectivity", *Proceeding of the National Academy of Sciences* **106** 2035

(2009).
[73] C. J. Honey, et al., "Can structure predict function in the human brain?", *Neuroimage* **52** 766 (2010) .
[74] R. Milo, et al., "Network motifs: simple building blocks of complex networks" *Science* **298**, 824 (2002).
[75] S. Mangan, and R. Milo, "Structure and function of the feed-forward loop network motif", *Proceeding of the National Academy of Sciences* **100** 11980 (2003).
[76] R. Milo, et al., "Superfamilies of evolved and designed networks" *Science* **303** (5663), 1538-1542 (2004).
[77] U. Alon, "Network motifs: theory and experimental approaches" *Nature Reviews Genetics* **8** (6), 450-461 (2007).
[78] Y. Adachi, et al., "Functional connectivity between anatomically unconnected areas is shaped by collective network-level effects in the macaque cortex", *Cerebral Cortex* **22**, 1586 (2012).
[79] K. Shen, et al., "Information processing architecture of functionally defined clusters in the macaque cortex", *Journal of Neuroscience* **32**, 17465 (2012).
[80] Ch. Li, "Function of neuronal network motifs", *Physical Review E* **78**, 037101 (2008).
[81] E. J. Friedman, et al., "Directed network motifs in alzheimer?s disease and mild cognitive impairment", *PLoS One* **10**, e0124453 (2015).
[82] O. Sporns, and R. Kötter, "Motifs in Brain Networks", *PLoS Biology* **2**, 1910 (2004).
[83] S. Sakata, et al., "Local design principles of mammalian cortical networks", *Neuroscience Research* **51**, 309 (2005).
[84] Y. Iturria-Medina, et al., "Studying the human brain anatomical network via diffusion-weighted MRI and Graph Theory", *Neuroimage* **40**, 1064 (2008).
[85] M. J. Brookes, et al., "A multi-layer network approach to MEG connectivity analysis", *Neuroimage* **132**, 425 (2016).
[86] M. De Domenico, S. Sasai, and A. Arenas, "Mapping multiplex hubs in human functional brain networks", *Frontiers in Neuroscience* **10**, 326 (2016).
[87] M. Kivela, and M. Porter, "Isomorphisms in multilayer networks", *arXiv preprint*, arXiv:1506.00508 (2015).
[88] L. Barrett, S.P. Henzi, and D. Lusseau, "Taking sociality seriously: the structure of multi-dimensional social networks as a source of information for individuals", *Philosophical Transactions of the Royal Society B* **367**, 2108-2118 (2012).
[89] P. Brodka, P. Kazienko, K. Musial, and K. Skibicki, "Analysis of neighbourhoods in multi-layered dynamic social networks" *International Journal of Computational Intelligence Systems* **5**, 582-596 (2012).
[90] R. Criado, J. Flores, A. Garcia del Amo, J. Gomez-Gardenes, and M. Romance, "A mathematical model for networks with structures in the mesoscale", *International Journal of Computer Mathematics* **89**, 291-309 (2011).
[91] F. Battiston, V. Nicosia, M. Chavez V. Latora, "Multilayer motif analysis of brain

networks", *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27 (4), 047404 (2017)

[92] J. A. Brown, and J. D. Van Horn , "Connected brains and minds ? the UMCD repository for brain connectivity matrices", *Neuroimage* **124**, 1238 (2016).

[93] J. D. Rudie, et al., "Altered functional and structural brain network organization in autism", *Neuroimage: clinical*, **2**, 79 (2013).

[94] J. D. Power, et al., "Functional network organization of the human brain", *Neuron* **72**, 665 (2011).

[95] D. Li, C. Zhou, "Organization of anti-phase synchronization pattern in neural networks: what are the key factors?", *Frontiers in Systems Neuroscience* **5**, 100 (2011).

[96] D. Vatansever, et al., "Default mode dynamics for global functional integration", *Journal of Neuroscience*, **35**, 15254 (2015).

[97] M. D. Fox et al., "The global signal and observed anticorrelated resting state brain networks", *Journal of Neurophysiology*, **101**, 3270 (2009).

[98] A. M. Kelly et al., "Competition between functional brain networks mediates behavioral variability", *Neuroimage*, **39**, 527 (2008).

[99] P. Tseng et al., "The critical role of phase difference in gamma oscillation within the temporoparietal network for binding visual working memory", *Scientific Reports* **6**, 32138 (2016).

[100] F. De Vico Fallani, J. Richiardi, M. Chavez, S. Achard, "Graph analysis of functional brain networks: practical issues in translational neuroscience", *Philosophical Transactions of the Royal Society B* **369**, 1653 (2014).

[101] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems", *Nature Review Neuroscience*, **10**, 1 (2009).

[102] T. Simas, et al., "An algebraic topological method for multimodal brain networks comparisons", *Frontiers in Psychology*, **6** (2015).

[103] J. S. Damoiseaux, et al., "Greater than the sum of its parts: a review of studies combining structural connectivity and resting-state functional connectivity", *Brain Structure and Function*, **213**, 525 (2009).

[104] A. T. Baria, et al., "Anatomical and functional assemblies of brain BOLD oscillations", *Journal of Neuroscience*, **31**, 7910 (2011).

[105] P. Skudlarski, et al., "Measuring brain connectivity: diffusion tensor imaging validates resting state temporal correlations", *Neuroimage* **43**, 554 (2008).

[106] A. Schlögel and G. Supp, *Progress in Brain Research* **159** 135, C. Neuper and W. Klimesh, Editors., (Elsevier), The Netherlands, 2006.

[107] A. Halu, S. Mukherjee, G. Bianconi, "Emergence of overlap in ensembles of spatial multiplexes and statistical mechanics of spatial interacting network ensembles", *Physical Review E*, **89**, 012806 (2014)

[108] D. Cellai, G. Bianconi, "Multiplex networks with heterogeneous activities of the nodes", *Physical Review E*, **93**, 032302 (2016)

[109] O. Sagarra, C. J. Perez Vicente, A. Diaz-Guilera, "Role of adjacency-matrix degen-

eracy in maximum-entropy-weighted network models", *Physical Review E* **92** (5), 052816 (2015)

[110] M. Diakonova, V. Nicosia, V. Latora, M. San Miguel, "Irreducibility of multilayer network dynamics: the case of the voter model", *New Journal of Physics*, **18** 023010 (2016)

[111] V. Nicosia, G. Bianconi, V. Latora, M. Barthelemy, "Growing multiplex networks *Physical Review Letters* **111** 058701 (2013)

[112] J, Y. Kim and K.-I. Goh, "Coevolution and correlated multiplexity in multiplex networks", *Physical Review Letters* **111**, 058702 (2013).

[113] N. Momeni, B. Fotouhi, "Growing multiplex networks with arbitrary number of layers", *Physical Review E*, **92**, 062812 (2015)

[114] PL Krapivsky, S Redner, F Leyvraz, "Connectivity of growing random networks" *Physical review letters* **85** (21), 4629 (2000)

[115] V. Nicosia, G. Bianconi, V. Latora, M. Barthelemy, "Nonlinear growth and condensation in multiplex networks", *Physical Review E* **90**, 042807 (2014)

[116] Y. Murase, J. Török, H.-H. Jo, K. Kaski, J. Kertész, "Multilayer weighted social network model", *Physical Review E*, **90**, 052810 (2014)

[117] S. John, "Social network analysis: A handbook" *London: SAGE Publications* (2000)

[118] JJ Ramasco, SN Dorogovtsev, R Pastor-Satorras, "Self-organization of collaboration networks" *Physical review E* **70** (3) 036106 (2004)

[119] A. Rapoport, "Spread of information through a population with socio-structural bias: I. Assumption of transitivity", *The bulletin of mathematical biophysics* **15** (4) 523–533 (1953)

[120] M. Newman, "Clustering and preferential attachment in growing networks", *Physical Review E* **64** (2) 025102 (2001)

[121] D. Lee, KI Goh, B Kahng, D. Kim, "Complete trails of coauthorship network evolution" *Physical Review E* **82** (2) 026112 (2010)

[122] P. Holme, B.J. Kim, "Growing scale-free networks with tunable clustering" *Physical review E* **65** (2) 026107 (2002)

[123] G. Bianconi, R.K. Darst, J. Iacovacci, S. Fortunato, "Triadic closure as a basic generating mechanism of communities in complex networks", *Physical Review E* **90** (4) 042806 (2014)

[124] M. Girvan, M. Newman, "Community structure in social and biological networks", *Proceedings of the national academy of sciences* **99** (12) 7821–7826 (2002)

[125] S. Fortunato, "Community detection in graphs", *Physics Reports* **486** (3) 75–174 (2010)

[126] M. De Domenico, A. Lancichinetti, A. Arenas, M. Rosvall, "Identifying modular flows on multilayer networks reveals highly overlapping organization in interconnected systems", *Physical Review X* **5** (1) 011027 (2015)

[127] J. Iacovacci, Z. Wu, G. Bianconi, "Mesoscopic structures reveal the network between

the layers of multiplex data sets", *Physical Review E* **92** (4), 042806 (2015)
[128] L. Danon, A. Diaz-Guilera, J. Duch, A. Arenas, "Comparing community structure identification", *Journal of Statistical Mechanics: Theory and Experiment* (09) P09008 (2005)
[129] M. Rosvall, C.T. Bergstrom, "Maps of random walks on complex networks reveal community structure" *Proceedings of the National Academy of Sciences*, **105** (4) 1118–1123 (2008)
[130] M.A.Serrano, M. Boguna, "Clustering in complex networks. I. General formalism" *Physical Review E*, **74** 056114 (2006)
[131] S.P. Borgatti, M.G. Everett, "Models of core/periphery structures", *Social networks* **21** (4), 375-395 (2000)
[132] G. Fagiolo, J. A. Reyes, S. Schiavo, "The Evolution of the World Trade Web: A Weighted-Network Analysis", *Journal of Evolutionary Economics*, **20**, 4, 479-514, (2009)
[133] J. P. Boyd, W. J. Fitzgerald, M. C. Mahutga, D. A. Smith, "Computing continuous core/periphery structures for social relations data with MINRES/SVD", *Social Networks*, **32** 125-137 (2010)
[134] F. Luo, B. Li, X.F. Wan, R.H. Scheuermann, "Core and periphery structures in protein interaction networks", *Bmc Bioinformatics* **10** (4), S8 (2009)
[135] M. P. Van den Heuvel, O. Sporns, "Rich-club organization of the human connectome", *Journal of Neuroscience* **31** (2011).
[136] F. Battiston, F. De Vico Fallani, M. Chavez, J. Guillon, V. Latora, "Rich cores in multiplex brain networks", in preparation (2017).
[137] P. Barucca, F. Lillo, "Disentangling bipartite and core-periphery structure in financial networks", *Chaos, Solitons & Fractals* **88**, 244-253 (2016)
[138] M.P. Rombach, M.A. Porter, J.H. Fowler, P.J. Mucha, "Core-periphery structure in networks", *SIAM Journal on Applied mathematics* **74** (1), 167-190 (2014)
[139] X. Zhang, T. Martin, M.E.J. Newman, "Identification of core-periphery structure in networks", *Physical Review E* **91** (3), 032803 (2015)
[140] A. Ma, R. J. Mondragon, "Rich-cores in networks", *Plos One*, **10**, (3) (2015).
[141] W.W. Zachary, "An information flow model for conflict and fission in small groups", *Journal of anthropological research* 452-473 (1977)
[142] A. Ma, R.J. Mondragon, V. Latora, "Anatomy of funded research in science", *Proceedings of the National Academy of Sciences* **112** (48), 14760-14765 (2015)
[143] T. Valles-Catala, F.A. Massucci, R. Guimera, M. Sales-Pardo, "Multilayer stochastic block models reveal the multilayer structure of complex networks", *Physical Review X* **6**, 011036 (2016).
[144] T.P. Peixoto, "Inferring the mesoscale structure of layered, edge-valued, and time-varying networks", *Physical Review E* **92** (4), 042807 (2015)
[145] GJ Baxter, D Cellai, SN Dorogovtsev, AV Goltsev, JFF Mendes, "A unified approach to percolation processes on multiplex networks", *Interconnected Networks*,

101-123 (2016)
[146] Z Wang, L Wang, A Szolnoki, M Perc, "Evolutionary games on multilayer networks: A colloquium", *The European Physical Journal B* **88**, 124 (2015)
[147] M Salehi, R Sharma, M Marzolla, M Magnani, P Siyari, D Montesi, "Spreading processes in multilayer networks", *IEEE Transactions on Network Science and Engineering* **2** (2), 65-83 (2015)
[148] M De Domenico, C Granell, MA Porter, A Arenas, "The physics of spreading processes in multilayer networks" *Nature Physics* **12**, 901-906 (2016)
[149] A. Barrat, M. Barthelemy, A. Vespignani, "Dynamical Processes on Complex Networks", *Cambridge University Press* New york, NY, USA (2008).
[150] J. D. Noh and H. Rieger, "Random Walks on Complex Networks", *Phys. Rev. Lett.* **92**, 118701 (2004).
[151] S.-J. Yang, "Exploring complex networks by walking on them", *Physical Review E* **71**, 016107 (2005).
[152] Z. Zhang, T. Shan and G. Chen, "Random Walks on Weighted Networks", *Physical Review E* **87**, 012112 (2013).
[153] Y.Lin and Z.Zhang, "Random walks in weighted networks with a perfect trap: An application of Laplacian spectra", *Physical Review E* **87**, 062140 (2013).
[154] H. Zhou, "Network landscape from a Brownian particles perspective", *Physical Review E* **67**, 041908 (2003).
[155] F. De Vico Fallani, V. Nicosia, V. Latora, M. Chavez, "Non-parametric resampling of random walks for spectral network clustering" *Physical Review E* **89**, 012802 (2014).
[156] V. Nicosia, M. De Domenico, V. Latora, "Characteristic exponents of complex networks", *Europhysics Letters* **106**, 58005 (2014).
[157] J. Gómez-Gardeñes, V. Latora, "Entropy Rate of Diffusion Processes on Complex Networks" *Physical Review E*, **78**, 065102(R) (2008).
[158] S. Lee, S. H. Yook and Y. Kim, "Centrality measure of complex networks using biased random walks", *The European Physical Journal B* **68**, 277–281 (2009).
[159] J.-C. Delvenne and A.-S. Libert, "Centrality measures and thermodynamic formalism for complex networks" *Physical Review E* **83**, 046117 (2011).
[160] V. Zlatić, A. Gabrielli and G. Caldarelli, "Topologically biased random walk and community finding in networks", *Physical Review E* **82**, 066109 (2010).
[161] A. Fronczak and P. Fronczak, "Biased random walks in complex networks: The role of local navigation rules", *Physicial Review E* **80**, 016107 (2009).
[162] R. Sinatra, J. Gómez-Gardenes, R. Lambiotte, V. Nicosia and V. Latora, "Maximal-entropy random walks in complex networks with limited information", *Physical Review E* **83**, 030103 (2011).
[163] A. Baronchelli and R. Pastor-Satorras, "Mean-field diffusive dynamics on weighted networks", *Physical Review E* **82**, 011111 (2010).
[164] M. Bonaventura, V. Nicosia and V. Latora, "Characteristic times of biased random

walks on complex networks", *Phys. Rev. E* **89**, 012803 (2014).
[165] F. Battiston, V. Nicosia, V. Latora "Efficient exploration of multiplex networks", *New Journal of Physics* **18** (4), 043035 (2016).
[166] M. De Domenico, A. Sole-Ribalta, S. Gomez, and A. Arenas. "Navigability of interconnected networks under random failures", *Proceeding of the National Academy of Sciences* **111** (23), 8351-8356 (2014).
[167] T. M. Cover and J. A. Thomas, "Elements of Information Theory", Wiley (1991).
[168] Z. Burda, J. Duda, J. M. Luck, and B. Waclaw, "Localisation of the maximal entropy random walk", *Physical Review Letters* **102**, 160602 (2009).
[169] M. Barthelemy, "Spatial networks", *Physics Reports* **499** 1-101 (2011).
[170] D. Lazer, A. Pentland, L. Adamic, S. Aral, A.-L. Barabasi, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, T. Jebara, G. King, M. Macy, D. Roy, M. V. Alstyne, "Social science: Computational social science", *Science* **323** (5915) 721-723 (2009)
[171] M. Jackson, "Social and Economic Networks", Princeton University Press, (2010)
[172] C Castellano, S Fortunato, V Loreto, "Statistical physics of social dynamics", *Reviews of modern physics* **81** (2), 591 (2009)
[173] F.-Y. Wu, "The potts model", *Review of Modern Physics* **54** (1) 235 (1982)
[174] P. Clifford, A. Sudbury, "A model for spatial conflict", *Biometrika* **60** (3) 581-588 (1973)
[175] R. A. Holley, T. M. Liggett, "Ergodic theorems for weakly interacting infinite systems and the voter model", *Annals of Probability* 643-663 (1975)
[176] S. Galam, "Majority rule, hierarchical structures, and democratic totalitarianism: A statistical approach", *Journal of Mathematical Psychology* **30** (4) 426-434 (1986)
[177] Vazquez, Federico and Eguíluz, Víctor M. and San Miguel, Maxi, *Physical Review Letters*, **100**, 108702 (2008)
[178] F.Vazquez,V.M.Eguiluz, "Analytical solution of the voter model on uncorrelated networks", *New Journal of Physics* **10** (6) 063011 (2008)
[179] R. Axelrod, "The dissemination of culture: a model with local convergence and global polarization", *Journal of Conflict Resolution*, **41,** (2), 203-226 (1997)
[180] S. Galam, "Minority opinion spreading in random geometry", *The European Physical Journal B* **25** (4) 403-406 (2002)
[181] J. Bryant, M. B. Oliver, "Media effects: Advances in theory and research", *Routledge*, (2008)
[182] T. V. Martins, M. Pineda, R. Toral, "Mass media and repulsive interactions in continuous-opinion dynamics", *Europhysics Letters* **91** (4) 48003 (2010)
[183] T. Carletti, D. Fanelli, S. Grolli, A. Guarino, "How to make an efficient propaganda", *Europhysics Letters* **74** (2) 222 (2006)
[184] W. Quattrociocchi, G. Caldarelli, A. Scala, "Opinion dynamics on interacting networks: media competition and social influence", *Scientific Reports* **4**. (2014)
[185] J. C. Gonzalez-Avella, M. G. Cosenza, V. M. Eguiluz, M. San Miguel, "Sponta-

neous ordering against an external field in nonequilibrium systems", *New Journal of Physics* **12** (1) 013010 (2008)

[186] J. C. Gonzalez-Avella, M. G. Cosenza, K. Tucci, "Nonequilibrium transition induced by mass media in a model for social influence", *Physical Review E* **72** (6) (2005)

[187] F. Battiston, A. Cairoli, V. Nicosia, A. Baule, V. Latora, "Interplay between consensus and coherence in a model of interacting opinions", *Physica D: Nonlinear Phenomena* **323**, 12-19 (2016)

[188] M. Diakonova, M. San Miguel and V. E. Eguíluz, *Physical Review E*, **89**, 062818 (2014)

[189] A. Chmiel, K. Sznajd-Weron, "Phase transitions in the q-voter model with noise on a duplex clique", *Physical Review E* **92** 052812 (2015).

[190] M Diakonova, V Nicosia, V Latora, M San Miguel, "Irreducibility of multilayer network dynamics: the case of the voter model" *New Journal of Physics* **18** (2), 023010 (2016)

[191] L. Onsager, "Crystal statistics. i. a two-dimensional model with an order-disorder transition", *Physical Review* **65** (3-4) 117 (1944)

[192] C. Doyle, S. Sreenivasan, B. Szymanski, G. Korniss, "Social consensus and tipping points with opinion inertia", *Physica A: Statistical Mechanics and its Applications* **443** 316-323 (2016)

[193] S. Galam, S. Moscovici, "Towards a theory of collective phenomena: consensus and attitude changes in groups", *European Journal of Social Psychology* **21** 49-74 (1991)

[194] S. Galam, "Rational group decision making: a random field Ising model at T=0", *Physica A: statistical mechanics and its applications* **238** 66-80 (1997)

[195] R. J. Glauber, "Time-dependent statistics of the ising model", *Journal of Mathematical Physics* **4** (2) (1963).

[196] P. Simon, F. Ricci-Tersenghi, "Coupled ising models with disorder", *Journal of Physics A: Mathematical and Theoretical* **33** (34) 5985 (2000)

[197] D. Baldassarri, , A. Gelman, "Partisans without Constraint: Political Polarization and Trends in American Public Opinion", *American Journal of Sociology* **114** (2), 408-446 (2008).

[198] F. Battiston, V. Nicosia, V. Latora, M. San MIguel, "Layered social influence promotes multiculturality in the Axelrod model", *Scientific Reports*, **7** 1809 (2017)

[199] B. Fredrik, "Ethnic and Group Boundaries" (Reissued Long Grove, IL: Waveland Press, 1998) (1969)

[200] R. Boyd, P.J. Richerson, "The Origin and Evolution of Cultures" (Oxford University Press, New York) (2005)

[201] L. Festinger, S. Schachter, K. Back, "Social Pressures in Informal Groups: a Study of Human Factors in Housing" (Stanford University Press, Palo Alto, CA) (1963)

[202] T.M. Liggett, "Interacting Particle Systems", (Springer, New York) (1985)

[203] G.C. Homans, "Sentiments and Activities" (Free Press, New York) (1962)

[204] J.M. McPherson, L. Smith-Lovin, "Homophily in voluntary organizations: Status distance and the composition of face-to-face groups" *American Sociological Review* 370-379 (1987)

[205] J.M. McPherson, L. Smith-Lovin, J. Cook "Birds of a feather: Homophily in social networks" *Annual Review of Sociology* **27** 415-444. (2001)

[206] C. Castellano, M. Marsili, A. Vespignani, "Nonequilibrium phase transition in a model for social influence" *Physical Review Letters* **85,** (16) 3536 (2000)

[207] K. Klemm, V.M. Eguiluz, R. Toral, M. San Miguel, "Nonequilibrium transitions in complex networks: A model of social interaction" *Physical Review E* **67,** 026120 (2003)

[208] K. Klemm, V.M. Eguiluz, R. Toral, M. San Miguel, "Global culture: A noise-induced transition in finite systems" *Physical Review E* **67,** (4), 045101 (2003)

[209] K. Klemm, V.M. Eguiluz, R. Toral, M. San Miguel, "Globalization, polarization and cultural drift" *Journal of Economic Dynamics and Control* **29** (1), 321-334 (2005)

[210] D. Centola, J.C. Gonzalez-Avella, V.M. Eguiluz, M. San Miguel, "Homophily, cultural drift and the co-evolution of cultural groups" *Journal of Conflict Resolution* **51** (6), 905-929 (2007)

[211] M. Sherif, C.I. Hovland, "Assimilation and Contrast Effects in Communication and Attitude Change" (Yale University Press, New Haven, CT) (1961)

[212] G. Deffuant, D. Neau, F. Amblard, G. Weisbuch, "Mixing beliefs among interacting agents" *Advances in Complex Systems* **3** (1), 87-98 (2000)

[213] G. Weisbuch, "Bounded confidence and social networks" *The European Physical Journal B* **38** 339-343 (2004)

[214] L. De Sanctis, T. Galla, "Effects of noise and confidence thresholds in nominal and metric Axelrod dynamics of social influence" *Physical Review E* **79** (4) 046108 (2009)

[215] L. Valori, F. Picciolo, A. Allansdottir, D. Garlaschelli, "Reconciling long-term cultural diversity and short-term collective social behavior" *Proceedings of the National Academy of Sciences*, **109** (4),1068-1073 (2011)

[216] A. Stivala, G. Robins, Y. Kashima, M. Kirley, "Ultrametric distribution of culture vectors in an extended Axelrod model of cultural dissemination", *Scientific Reports* **4** (2014)

[217] A.I. Babeanu, L. Talman, G. Garlaschelli, "Universal properties of culture: evidence for mixed rationalities in preference formation", *arXiv preprint:1506.01634* (2016)

[218] A. Flache, M. Macy, "What sustains cultural dversity and what undermines it? Axelrod and beyond" *ArXiv preprint: 0604201* (2006)

[219] A. Flache, M. Macy, "Local convergence and global diversity: From interpersonal to social influence" *Journal of Conflict Resolution* **55** 6 968-993 (2011)

[220] M. Granovetter, "Threshold Models of Collective Behavior" *American Journal of Sociology* **83** 6, 1420-1433 (1978)
[221] D.J. Watts, "A simple model of global cascades on random networks", *Proceedings of the National Academy of Sciences* **99** (9), 5766-5771 (2002)
[222] D. Centola, "The spread of behavior in an online social network experiment", *Science* 329, 1194 (2010)
[223] D. Centola, V.M. Eguiluz, M. Macy, "Cascade dynamics of complex propagation", *Physica A* 374, 449-456 (2007)
[224] A. Flache, M. Macy, "Small worlds and cultural prolarization", *The Journal of Mathematical Sociology* **35** 146 (2011)
[225] M. Mas, A. Flache, J.A. Kitts, "Cultural integration and differentiation in groups and organizations" (Springer) (2014)
[226] D.M. Abrams, S.H. Strogatz, "Chimera states for coupled oscillators" *Physical Review Letters* **93** (17) 174102 (2004)
[227] M.R. Tinsley, S. Nkomo, K. Showalter, "Chimera and phase-cluster states in populations of coupled chemical oscillators", *Nature Physics* **8** 662 (2012)
[228] J.C. Gonzalez-Avella, M.G. Cosenza, M. San Miguel, "Localized coherence in two interacting populations of social agents" *Physica A* **399** 24-30 (2014)
[229] F. Battiston, M. Perc, V. Latora, "Determinants of public cooperation in multiplex networks", *New Journal of Physics* **19** 071002 (2017)
[230] M.A. Nowak, R. Highfield, "SuperCooperators: Altruism, Evolution, and Why We Need Each Other to Succeed", Free Press, New York (2011)
[231] S.B. Hrdy, "Mothers and Others: The Evolutionary Origins of Mutual Understanding" Harvard University Press, Cambridge, MA, (2011)
[232] K. Sigmund, "Games of Life: Exploration in Ecology, Evolution and Behavior", Oxford University Press, (1993)
[233] J. W. Weibull, "Evolutionary Game Theory", MIT Press, Cambridge, MA (1995)
[234] J. Hofbauer, K. Sigmund, "Evolutionary Games and Population Dynamics", Cambridge University Press, UK (1998)
[235] M.A. Nowak, "Evolutionary Dynamics", Harvard University Press, Cambridge, MA (2006)
[236] K. Sigmund, "The Calculus of Selfishness", Princeton University Press, NJ (2010)
[237] M. Perc, J. Gómez-Gardeñes, A. Szolnoki, L. M. Floría, Y. Moreno, "Evolutionary dynamics of group interactions on structured populations: a review", *Journal of the Royal Society Interface*, **10**, 20120997 (2013)
[238] G. Hardin, "The Tragedy of the Commons", *Science*, **162**, 1243-1248 (1968)
[239] M.A. Nowak, "Five Rules for the Evolution of Cooperation", *Science*, **314**, 1560-1563 (2006)
[240] D.A. Rand, M. A. Nowak, "Human cooperation", *Trends in Cognitive Sciences*, **17**, 413-425 (2013)
[241] M. A. Nowak, R. M. May, "Evolutionary Games and Spatial Chaos", *Nature*, **359**,

826-829 (1992)
[242] György Szabó and Gábor Fáth, "Evolutionary games on graphs", *Physics Reports*, **446**, 97-216 (2007)
[243] M. Perc, A. Szolnoki, "Coevolutionary games – a mini review", *BioSystems*, **99**, 109-125 (2010)
[244] J. M. Pacheco, V. V. Vasconcelos, F. C. Santos, "Climate change governance, cooperation and self-organization", *Physics of Life Reviews*, **11**, 573-586 (2014)
[245] Z. Wang, A. Szolnoki, M. Perc, "Evolution of public cooperation on interdependent networks: The impact of biased utility functions", *Europhysics Letters*, **97**, 48001 (2012)
[246] J. Gómez-Gardeñes, I. Reinares, A. Arenas, L. M. Floría, "Evolution of Cooperation in Multiplex Networks", *Scientific Reports*, **2**, 620 (2012)
[247] J. Gómez-Gardeñes, C. Gracia-Lázaro, L. M. Floría, Y. Moreno, "Evolutionary dynamics on interdependent populations",*Physical Review E*, **86**, 056113 (2012)
[248] B. Wang, X. Chen, L. Wang, "Probabilistic interconnection between interdependent networks promotes cooperation in the public goods game", *Journal of Statistical Mechanics: Theory and Experiments*, **2012**, P11017 (2012)
[249] Z. Wang, A. Szolnoki, M. Perc, "Interdependent network reciprocity in evolutionary games", *Scientific Reports*, **3**, 1183 (2013)
[250] A. Szolnoki, M. Perc, "Information sharing promotes prosocial behaviour", *New Journal of Physics*, **15**, 053010 (2013)
[251] Z. Wang, A. Szolnoki, M. Perc, "Self-organization towards optimally interdependent networks by means of coevolution", *New Journal of Physics*, **16**, 033041 (2014)
[252] G. Szabó, C. Tőke, "Evolutionary prisoner's dilemma game on a square lattice", *Physical Review E*, **58**, 69-73, (1998)
[253] A. Szolnoki, M. Perc, G. Szabó, "Topology-independent impact of noise on cooperation in spatial public goods games", *Physical Review E*, **80**, 056109, (2009)
[254] M. A. Javarone, F. Battiston, "The role of noise in the spatial public goods game", *Journal of Statistical Mechanics: Theory and Experiment*, **7**, 073404 (2016)
[255] V. Nicosia, F. Battiston, "Metrics And Models for MULTilayer networks", *https* : *//github.com/fede7j/mammult* (2016)
[256] M. Starnini, A. Baronchelli, R. Pastor-Satorras, "Temporal correlations in social multiplex networks", arXiv preprint arXiv:1606.06626 (2016)
[257] M.E.J. Newman, A. Clauset, "Structure and inference in annotated networks", *Nature Communications* **7** 11863 (2016)
[258] D. Hric, T.P. Peixoto, S. Fortunato, "Network structure, metadata, and the prediction of missing nodes and annotations", *Physical Review X* **6** (3), 031038 (2016)